DESIGN, SYNTHESIS, AND BIOLOGICAL ACTIVITIES OF SMALL MOLECULES THAT TARGET MYOTONIC DYSTROPHY

BY

LIEN NGUYEN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Steven C. Zimmerman, Chair and Director of Research
Professor Paul J. Hergenrother
Professor John A. Katzenellenbogen
Professor Eric Oldfield

ABSTRACT

Myotonic dystrophy (DM) is a triple-repeat expansion, multi-systemic disease that affects one in eight thousand people worldwide. The cause of the disease is a progressive, abnormal expansion of CTG repeats (CUG^{exp}) in the 3'-UTR of the *DMPK* gene (DM1) and CCTG repeats ($CCUG^{exp}$) in the intron 1 of the *ZNP9* gene (DM2). The sequestration of muscleblind-like proteins (MBNL) by CUG^{exp} or $CCUG^{exp}$ causes splicing defects in more than 100 pre-mRNAs, resulting in various disease phenotypes. As such, therapeutic development for DM has mainly focused on agents targeting the CUG^{exp}/$CCUG^{exp}$ -MBNL1 interaction. This dissertation focuses on the development of rationally designed small molecules that target CUG^{exp} and $CCUG^{exp}$, including their synthesis and studies of their biological activity.

The background of DM with a focus on its molecular mechanism and various therapeutic approaches are reviewed in Chapter 1. Chapter 2 includes the story of how simple ligands targeting CUG^{exp} have been developed and investigations of the biological activity of acridine-based and bisamidinium-based ligands synthesized by others in our group. Chapter 3 focuses on the design, synthesis, and biological activity of bisamidinium-based ligands that target $CCUG^{exp}$.

Because other toxic pathways, including microRNA dysregulation in DM1 heart tissue and the production of polypeptides via repeat-associated non-ATG translation, are induced by CUG^{exp}, recent efforts on DM1 therapeutic approaches

have moved beyond just preventing the formation of CUG^{exp}-MBNL1 complex and further focused on regulating the level of toxic CUG^{exp}. Thus, Chapter 4 discusses a multi-target approach for DM1 in which a ligand with a RNA-cleaving unit can bind both CTG^{exp} and CUG^{exp} and regulate the level of CUG^{exp}.

I dedicate this work to my parents, my younger brother, and my close friends…

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor – Prof. Steven C. Zimmerman who has provided guidance, encouragement, and resources throughout my Ph.D. study. He is one of two people I admire the most, who affected my career choice and helped me to realize what I really want to do. With his help and support, I have been able to improve myself and feel more confident. He is also a person who helps me to establish a high standard in my work and remind me never forget to question myself and criticize my own work. I would like to offer my special thanks to my doctoral committee members Prof. John A. Katzenellenbogen, Prof. Paul J. Hegenrother, and Prof. Eric Oldfield for their valuable suggestions and guidance over the last four and a half years. I want to gratefully acknowledge Prof. Jim Lisy and Prof. Alexander Scheeline for their encouragement and caring. I would also like to thank Prof. Auinash Kalsotra for his help and suggestions.

I especially thank Long Luu and JuYeon Lee who are my great lab-mates, collaborators, and friends. My special thanks are extended to Brenda Andrade, Li Ying, Dr. Yugang Bai, Dr. Shampa Samanta, and Dr. Dawn Ernenwein for many great discussion and for their advice. I would like to especially thank my great mentors, Dr. Chun-Ho Wong, Prof. Si-Kyung Yang, and Dr. Stacie Richardson for teaching me science and for their moral support. I would like to thank Julio F. Serrano and Iti Kapoor for their collaboration and support. It has been wonderful to be in the DM1 subgroup with them. I thank other members of Prof. Zimmerman's group for their support and creating a great research atmosphere.

I would like to express my special thanks to my great Vietnamese friends, Huong Luu's

family, Linh-Duc Nguyen's family, Ly Luu's family, Han-Duc Anh's family, and other friends. With their help and caring, I am able to keep my life balanced. I especially thank my closest friends, Hien Nguyen and Tam Do, my cousin Bac Nguyen for their encouragement throughout years and for always staying by my side.

Lastly, I would like to express my deep gratitude to my parents and my great appreciation to my younger brother for their understanding, encouragement, and support. Thanks for being with me in all circumstances, for helping me overcome all difficulties in my life, and for giving me motivation.

TABLE OF CONTENTS

Chapter 1

Introduction to Myotonic Dystrophy

1.1. Disease pathogenesis of Myotonic Dystrophy

Myotonic dystrophy is a neuro-muscular, multisystem disease, affecting 1 in 8000 people worldwide.[1] There are two types of myotonic dystrophy: myotonic dystrophy type 1 (DM1) and myotonic dystrophy type 2 (DM2).[1-2] DM1 is caused by an expansion of CTG repeats from a normal level of 35–50 to several thousands repeats in the 3'-untranslated region of the *DMPK* gene on chromosome *19q13*.[2] DM2 occurs when there is a 75- to 11000-CCTG repeat expansion in the first intron of the *ZNF9* gene on chromosome *3q21* (Figure 1.1)[1].[1-4] The length of the expanded repeats is directly related to severity and onset of the disease.[5]

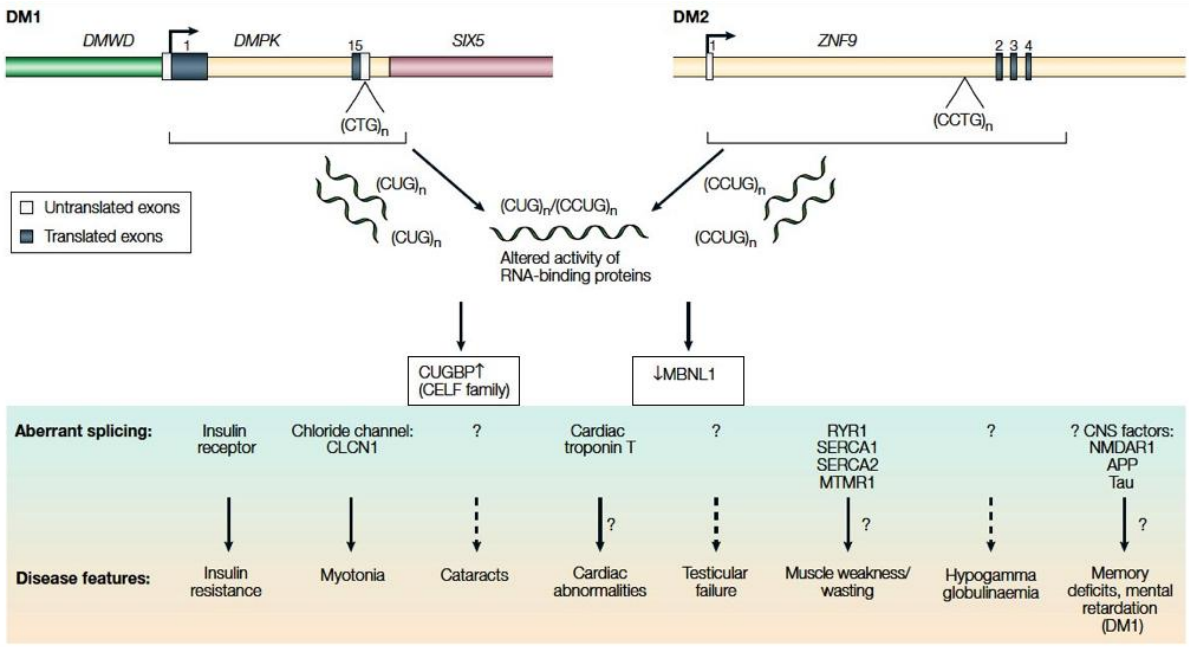


Figure 1.1. Mutated genes causing DM and early understanding in disease pathogenesis. The imbalance in MBNL and CELF family proteins results in missplicing of pre-mRNA directly correlating with disease phenotypes.

[1] The figure 1.1 was adapted from the following reference:
Gatchel, J. R.; Zoghbi, H. Y. *Nat. Rev. Genet.* **2005**, *6*, 743–755.

It has been shown that the RNA transcripts from CTG and CCTG repeats are toxic and the causative agent of the disease. The RNA transcripts exert this effect by sequestering the muscleblind-like (MBNL) protein family, including the alternate splicing regulator MBNL1.[4] Additionally, there is an increase in the level of CUG-binding protein 1 (CUG-BP1), another splicing regulator; the process by which this happens is currently unknown. The decrease in MBNL1 levels and elevation of CUG-BP1 levels are hypothesized to lead to splicing defects of a large set of pre-mRNAs, which subsequently lead to DM symptoms of myotonia, muscle weakness, cataracts, hypertrophy, and insulin resistance (Figure 1.1).[5]

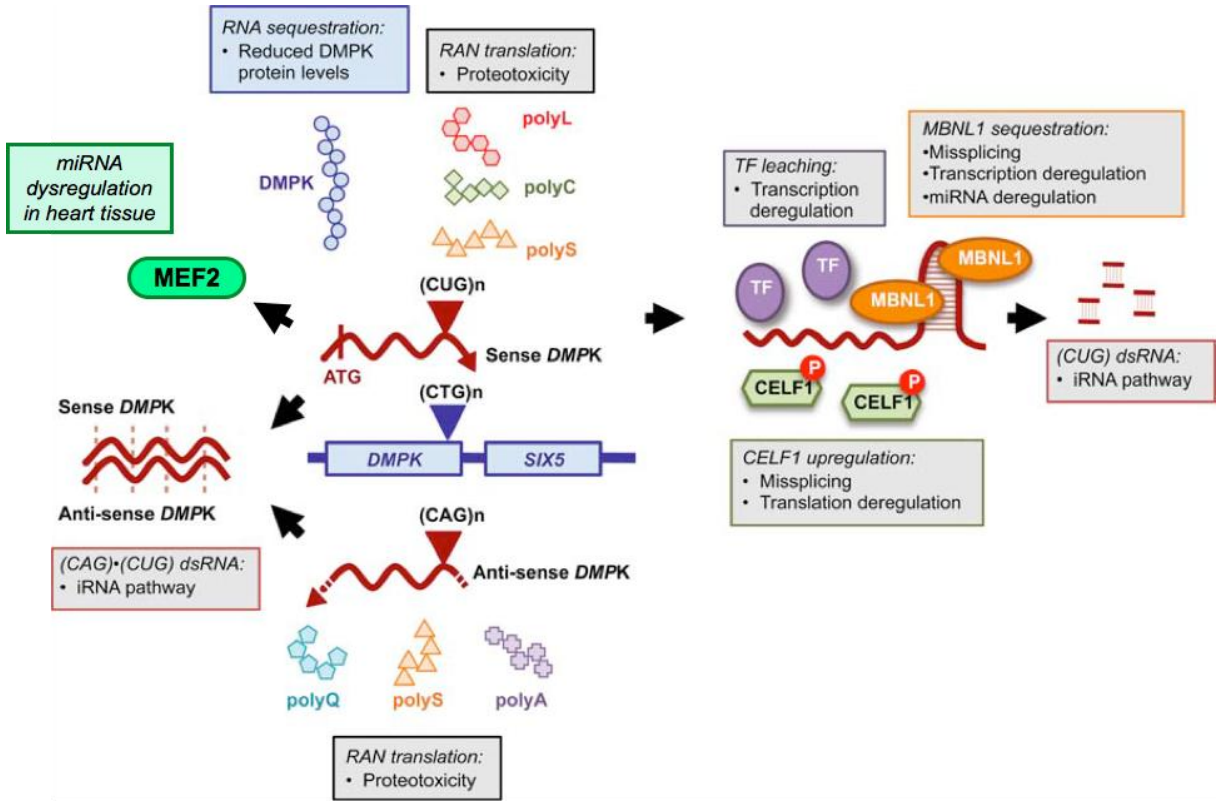


Figure 1.2. DM1 pathogenesis. $(CTG{\cdot}CAG)_n$ undergoes bi-directional transcription, generating $(CUG)_n$ and $(CAG)_n$ RNA. These two transcripts are translated to produce polypeptides that are potentially toxic. The CUG transcript sequesters MBNL proteins, preventing them from the normal functions and upregulates CELF1 (CUG-BP1) protein. The change in levels of MBNL and CELF1 proteins results in splicing defects of more than 100 pre-mRNAs, causing disease phenotypes at different levels. CUG repeats also interfere with iRNA and miRNA pathways.

In addition, the CTG repeats were shown to undergo bi-directional transcription resulting in two transcripts, namely CUG and CAG repeats (CUG^{exp} and CAG^{exp}). These transcripts are translated via repeat-associated non-ATG (RAN) translation to produce homopeptides, some of which were found to be toxic in other neurodegenerative diseases.[6] Thus, Ranum and colleagues reported the existence of polypeptides in disease-relevant tissues of DM1 patients.[7-9] Recently, Kalsotra and co-workers reported that the CUG repeat dysregulates the translation of Mef2 protein, causing the microRNA dysregulation in DM1 heart tissues (Figure 1.2)[2].[10,11] Although the understanding of the molecular mechanism of the disease has expanded, there currently is no treatment for DM. The following sections will discuss DM1 therapeutic developments.

1.2. Targeting DM1 at protein and DNA levels

DM therapeutics have been developed based on the above understanding of the molecular

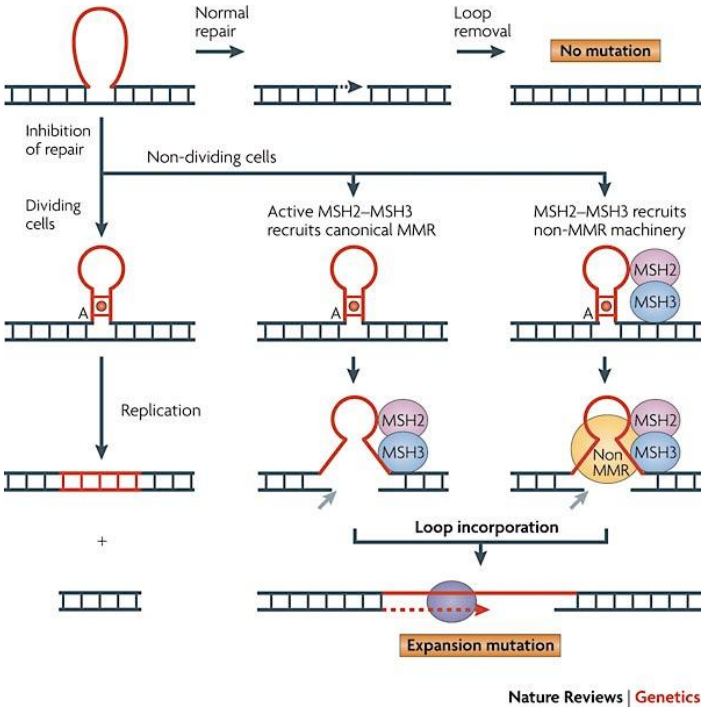


Figure 1.3. Proposed repeat expansion mechanism occurred during DNA replication and repair.

[2] The figure 1.2 was adapted from the following reference:
Pearson, C. E. *PLoS Genet.* **2011**, *7*, e1002018.

mechanism of the disease. Thus, supported by the discovery of disease foci, i.e., the MBNL1 sequestration by CUG^{exp} and $CCUG^{exp}$, and the function of MBNL1 protein in cells, an early potential therapeutic approach for DM was to increase the MBNL1 level in disease cells. For example, it was shown that overexpressing MBNL1 could suppress the CUG^{exp}-induced toxicity in DM1 *Drosophila* and mouse models.[13-15] Although the protein approach has a practical limitation, it supports the hypothesis that MBNL1 sequestration is an important event, thus encouraging scientists to pursue the CUG^{exp}-targeting approach for disease treatment.

Targeting the disease at the DNA level to induce a contraction of CTG·CAG or CCTG·CAGG repeats is considered the ultimate treatment for DM. Thus, a number of studies have been carried out to understand repeat expansion mechanism. Experiments using simple repeat expansion models in yeast, *E. coli*, mammalian cells, and mouse models have showed some consistency in the threshold of the repeat number required for the abnormal repeat expansion occurring in DM patients.[16-18] Importantly, it was shown that the repeat expansion could happen during replication, DNA repair, DNA recombination, and transcription with the formation of $(CTG)_n$ and $(CAG)_n$-loop regions as a key event.[12] Whereas the replication factors (e.g. helicases, polymerases) could contribute significantly to repeat instability at the early stage of cells, repair proteins (e.g. MSH2 and MSH3) and the factors that are key components in the remaining processes are believed to play important roles in $(CTG{\cdot}CAG)_n$ expansion at the later stage when replication is limited (Figure 1.3)[3].[12] Despite a significant effort to develop model systems, a precise understanding of how $(CTG{\cdot}CAG)_n$ undergoes expansion is lacking, which limits the development of DNA-targeting therapeutics. In fact, several known anticancer agents

[3] The figure 1.3 was adapted from the following reference:
McMurray, C. T. *Nat. Rev. Genet.* **2010**, *11*, 786–799.

were tested in DM patient cells showing random effects on $(CTG{\cdot}CAG)_n$ instability including no change, repeat expansion, and contraction (Table 1.1)[4].[18] Although tested compounds possess off-target activities causing cellular toxicity, the study supports an approach of using small molecule to contraction of $(CTG{\cdot}CAG)_n$ to its normal length.

Table 1.1. Effects of chemicals on repeat instability

Chemical	Effect on cell/DNA metabolism	Cell type tested	Effect on the dynamics of expanded repeats
Aphidicolin	Inhibition of DNA polymerase α, δ and ε	DM1 foetal primary fibroblasts	Increase in repeat expansion rate
araC	Inhibition of DNA polymerase α, δ and ε	*Dmt*-D mouse kidney cells	Reduction in repeat expansion rate
Aspirin	Accumulation of cells in G0/G1, suppression of the mutator phenotype associated with HNPCC	*Dmt*-D mouse kidney cells	Reduction in repeat expansion rate
5-Azacytidine	Induction of DNA hypomethylation	*Dmt*-D mouse kidney cells	Reduction in repeat expansion rate
5-Aza-2-deoxycytidine	Induction of DNA hypomethylation	CHO transgenic model	Increase in frequency of large deletions
5-Aza-2-deoxycytidine	Induction of DNA hypomethylation	DM1 fibroblasts	Increased frequency of large length changes
Caffeine	Inhibition of G2/M DNA damage checkpoint	*Dmt*-D mouse kidney cells	Increase in repeat expansion rate
Doxorubicin	Intercalation into DNA and inhibition of DNA synthesis	DM1 LBCL	Induction of large deletions
Emetine	Inhibition of Okazaki fragments synthesis	DM1 foetal primary fibroblasts	Increase in repeat expansion rate
EMS	DNA alkylation	DM1 LBCL	Accumulation of small alleles
Ethidium bromide	Intercalation into DNA, induction of oxidative stress	*Dmt*-D mouse kidney cells	Reduction in repeat expansion rate
Hydrogen peroxide	Induction of oxidative stress	*Dmt*-D mouse kidney cells	Reduction in repeat expansion rate
Mimosine	Inhibition of replication initiation	DM1 foetal primary fibroblasts	Minor effect
Mitomycin C	Induction of DNA intra-strand crosslinks	DM1 LBCL	Increase in repeat expansion rate
Mitomycin C	Induction of DNA intra-strand crosslinks	DM1 LBCL	Reduction in repeat expansion rate
Mitoxantrone	Inhibition of topoisomerase II	DM1 LBCL	Induction of large deletions
Novobiocin	Inhibition of topoisomerase II	*Dmt*-D mouse kidney cells	Minor effect
Rhodamine-6G	Induction of oxidative stress	*Dmt*-D mouse kidney	Reduction in repeat

[4] The table 1.1 was adapted from the following reference:
Gomes-Pereira, M.; Monckton, D. G. *Mutat. Res. - Fund. Mol. M.* **2006**, *598*, 15–34.

1.3. Targeting DM at the RNA level

Development of DM therapeutics has been mainly focused on targeting CUG^{exp} or $CCUG^{exp}$ to increase the level of MBNL1 in disease-model cells and animals. This focus is supported by the studies in which overexpression of MBNL1 in DM1 animal models could rescue several disease symptoms (discussed above). In addition, because CUG and CCUG repeats are located on non-coding regions, targeting these repeats is believed to minimize unwanted effects. Thus, the structures of CUG and CCUG repeats and the binding mode of MBNL1 and the repeats were investigated. X-ray studies on simplified sequences containing CUG or CCUG repeats showed that the repeat adopts A-form duplex with UU or CU mismatches separated by GC base pairs, respectively.[20,21] In addition, the CCUG transcript was hypothesized to have a slipped form that contains adjacent UU and CC mismatches.

It was reported that MBNL1 has a preference for binding pyrimidine mismatches, but binding affinities, albeit somewhat lower, were also observed for other mismatches. The binding of MBNL1 to CUG^{exp} requires 2 UU mismatches and 4 GC base pairs with an involvement of four zinc fingers, sharing similarities with its pre-mRNA substrate cTNT.[22] Thus, design of inhibitors of this interaction needs to consider the similarity of CUG^{exp} and the substrates carefully to minimize nonspecific binding. The MBNL1-$(CUG)_n$ binding was also studied using

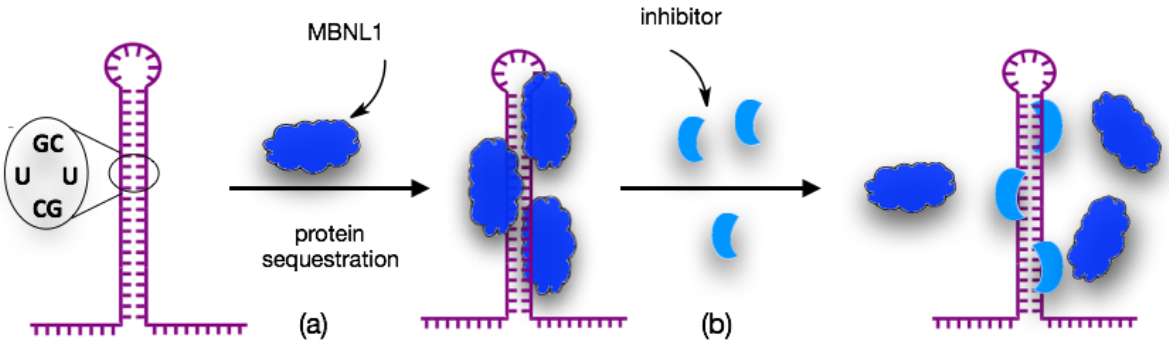


Figure 1.4. Interrupting MBNL1-CUG^{exp} interaction by small-molecule inhibitors. (a) CUG^{exp} forms a stable hairpin containing UU mismatches, which sequesters MBNL1 protein. (b) Inhibitors compete with MBNL1 to bind to CUG^{exp}, displacing MBNL1.

single molecule techniques, showing that the MBNL1 binding stoichiometry is CUG^{exp}-length dependent.[23] Specifically, $(CUG)_4$ is bound by one MBNL1 , whereas three MBNL1 molecules bind $(CUG)_{12}$ simultaneously. Although the binding of MBNL1-CUG^{exp} has not been studied in cells or *in vivo* at a molecular level, the co-localization of MBNL1 and CUG^{exp} in DM1 patient cells and DM1 model cells (i.e., HeLa transfected with a $(CTG)_n$ containing plasmid) provides strong evidence of the protein sequestration. Recent studies showed that the MBNL1-CUG^{exp} interaction could be modified by other factors in cells.[24,25]

CUG^{exp}, the causative agent of DM1, has been targeted using antisense and non-antisense approaches. In particular, antisense oligonucleotides or small molecules that selectively bind to CUG^{exp} can inhibit its interaction with MBNL1 allowing the liberated MBNL1 to function normally (Figure 1.4). Several oligonucleotides were employed to disrupt the MBNL1-CUG^{exp} interaction, and found to reverse the DM1 symptoms in *Drosophila* and mouse models.[8,26,27] The modified versions of oligonucleotides not only inhibit the formation of MBNL1-CUG^{exp} complexes but also regulate the level of toxic RNA via an RNase H-mediated mechanism.[27] This approach is believed to terminate other toxic pathways caused by CUG^{exp} such as RAN translation and miRNA dysregulation. Although the antisense treatment requires an intramuscular injection or electroporation to increase the cellular uptake, an investment in the

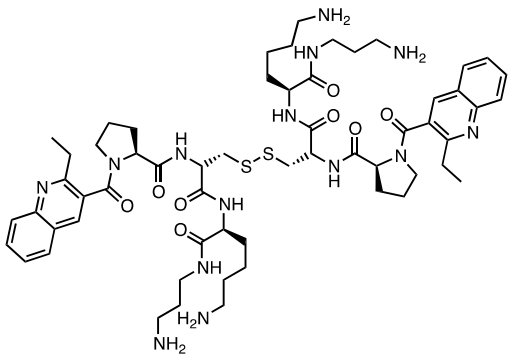


Figure 1.5. Peptide-based ligand developed by the Miller group.

antisense approach by Ionis (formally ISIS) Pharmaceuticals has lead to phase II clinical trials as of early 2015.

Whereas an effective method for the delivery of antisense agents may be still required for better efficacy, small molecules offer an alternative for DM1 treatment with their better cell permeability. Thus, we and other groups have focused on the development of small-molecule inhibitors based on multiple strategies including rational design, library screening, and combinatorial chemistry for past 8 years. This chapter covers the work on CUG^{exp}-targeting small molecules reported by other groups; our work will be discussed in Chapter 2.

In 2008, Miller and co-workers reported the use of resin-bound dynamic combinational chemistry to prepare the first examples of small molecule inhibitors of MBNL1-$(CUG)_n$, which exhibited *in vitro* K_i values in the low micromolar range (Figure 1.5).[28] Berglund and co-workers discovered that pentamidine and neomycin (Figure 1.6) could act as potential DM1 therapeutics and tested these compounds in the first cellular experiments in 2009. Pentamidine was tested further in a DM1 mouse model and showed partial rescue of splicing defects of *Clc-1* and *Serca1*. However, the further application was limited by the high toxicity of this agent.[29] A structure-activity relationship study was employed to develop pentamidine analogs with lower cytotoxicity and similar or even better bioactivities. Although none of pentamidine analogs

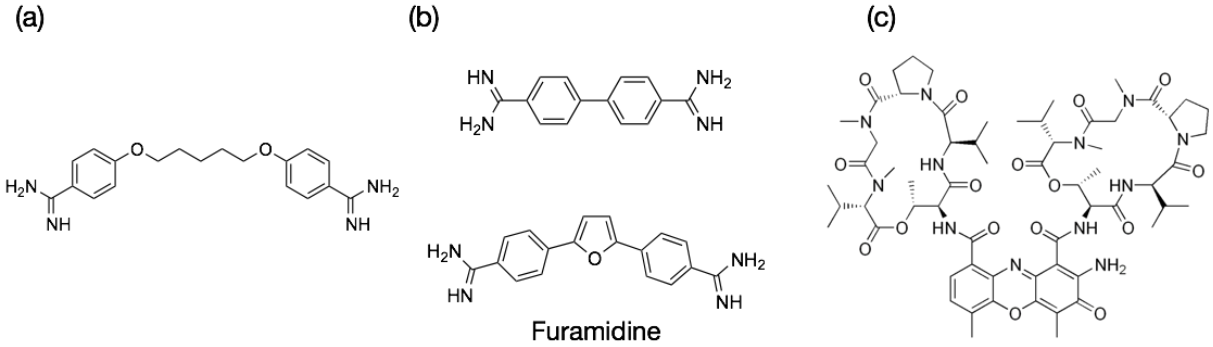


Figure 1.6. Potential DM1 therapeutic ligands found by Berglund group. (a) Pentamdine. (b) Diamidines. (c) Actinomycin D.

match the criteria, the Berglund group recently reported two diamidines (Figure 1.6b), known DNA group binders, as effective agents that were shown to fully rescue splicing defects of *IR* and *cTNT* pre-mRNA in a DM1 cell model. Interestingly, furamidine, one in two above diamidines, fully reversed *Clcn1* pre-mRNA missplicing in a DM1 mouse model.[29]

Very recently, actinomycin D (Figure 1.6c), an FDA-approved chemotherapeutic that binds to GC-rich DNA, was reported by the same group to suppress the level of toxic CUG repeats in the nanomolar concentration range.[30] This finding supports the therapeutic approach of targeting

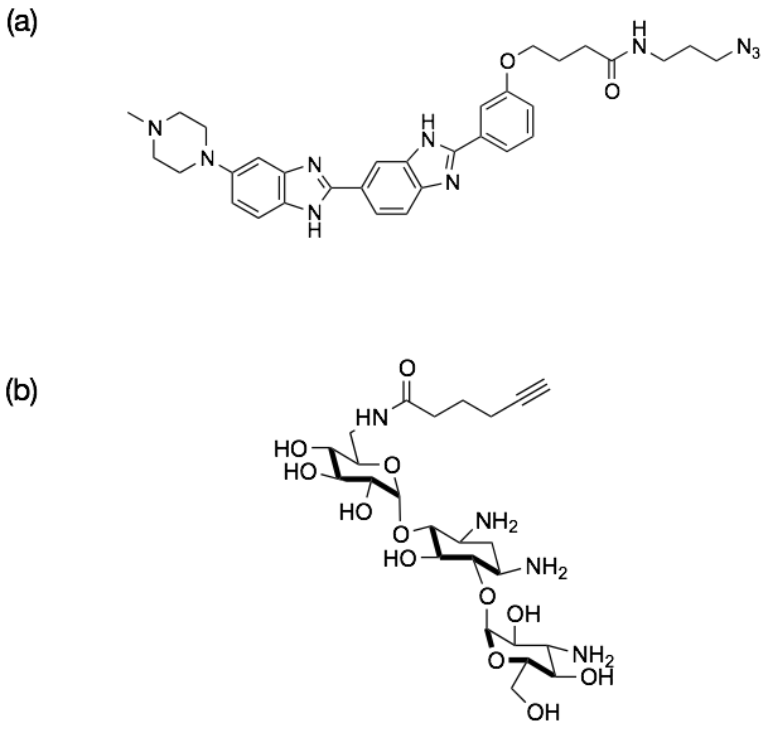


Figure 1.7. Compounds developed by the Disney group. (a) Hoechst 33258 derivative containing an azido handle. (b) 6'-*N*-5-Hexynoate kanamycin A derivative containing an alkyne group.

the production of CUG repeats in disease cells, which is also focused in our group (see Chapter 4). This approach not only prevents the formation of MBNL1-CUG^{exp} complexes but also eliminates other toxic pathways induced by CUG^{exp}.

Disney and co-workers have developed the other set of CUG^{exp}-targeting compounds based on Hoechst and kanamycin A (Figure 1.7), known DNA and RNA binders, respectively.[32,33] One of the most potent Hoechst-based compounds partially rescued the splicing defects of *cTNT* pre-RNA in a DM1 cell culture and *Clc-1* and *Serca1* in a DM1 mouse model.[34] Later, a series of assembled compounds containing multiple Hoechst molecules on a peptide scaffold showed a remarkable improvement in efficacy in DM1 model cells.[35] Regarding aminoglycoside-based compound-development, they improved the potency and cellular uptake of kanamycin A-peptoid assemblies by conjugating them with D-Arg9, which can act as molecular transporter.[36] To

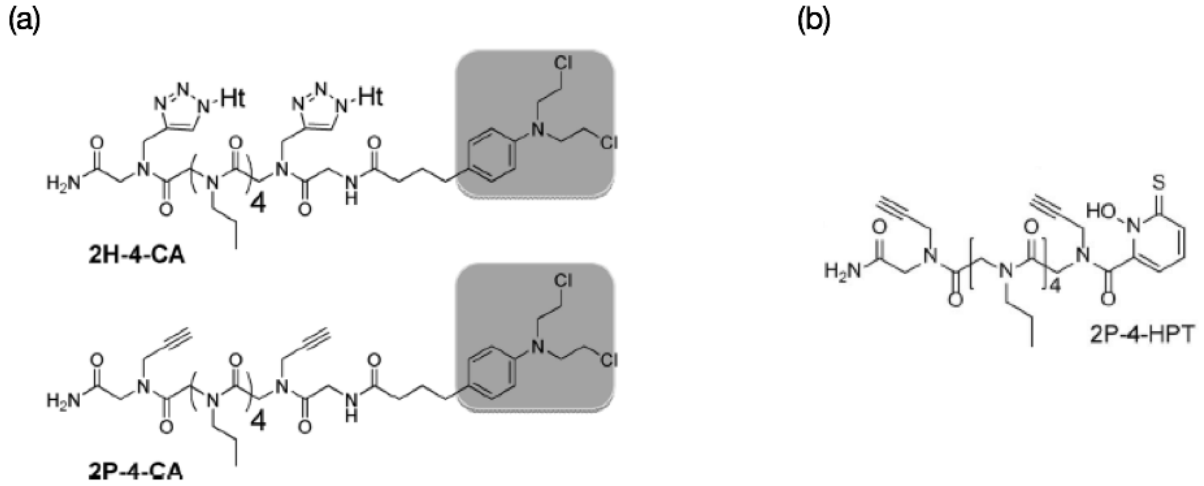


Figure 1.8. Covalent molecules (a) and a scaffold containing photo-induced RNA cleaving group for conjugating with azido containing Hoechst 33258 (b).

improve the potency of small molecules without a significant increase in molecular weight, Disney group developed a small molecule with a reactive group (Figure 1.8a)[5], which can form a covalent complex with the target CUG^{exp}. The covalent inhibitor with 2500-fold improvement in cellular activates was also used in a pull-down assay to determine its main targets in cells.[37] Because CUG^{exp} can cause other toxic pathways beside MBNL1 sequestration, Disney and co-

[5] The figure 1.8 was adapted from the following reference:
Guan, L.; Disney, M. D. *Angew. Chem. Int. Ed.* **2013**, *52*, 10010–10013.

workers synthesized a bifunctional small molecule that targets CUG^{exp} and mediates photo-induced RNA cleavage (Figure 1.8b).[38]

The development of compounds targeting $CCUG^{exp}$ for DM2 treatment has been transformed from DM1 targeting strategies because of similarities in the molecular mechanisms of DM1 and DM2. In 2009, a screening assay by the Disney group showed kanamycin A (Figure 1.7b) possessed a binding affinity toward $CCUG^{exp}$.[33] By conjugating kanamycin A on a peptoid backbone, the resulting assemblies inhibited MBNL1-$CCUG^{exp}$ at a nanomolar concentration. However, the huge increase in molecular weight and size of the compound could be problematic for water solubility and cellular uptake. Thus, they recently reported an approach of using $CCUG^{exp}$ as a catalyst to assemble its own potent inhibitors *in situ*.[39]

1.4. Conclusion

Myotonic dystrophy is a model example of repeat diseases, having significant effects on human health quality and subsequent generation. Although discovered almost 100 years ago and undergoing an expansion in understanding of the disease pathogenesis within the past 20 years, there is still no cure for DM. Due to an urgent need of DM therapeutics, multiple groups worldwide have focused on different approaches to treat DM. One of the most promising candidates for DM1 therapeutics is an antisense oligonucleotide that is under a clinical phase II. Another alternative of using an antisense approach to overcome the limitation in cellular uptake and for developing an oral drug is using a small molecule approach. In this Chapter, development of small molecules or small molecule-based compounds by the Berglund and Disney groups was focused with the starting point from the hits found from library screenings. In Chapter 2, our work during past 8 years on rational design of small molecules based on X-ray structures of CUG^{exp} and $CCUG^{exp}$ and their biological activities will be discussed.

1.5. Reference

(1) Mulders, S. A. M.; van Engelen, B. G. M.; Wieringa, B.; Wansink, D. G. *Hum. Mol. Genet.* **2010**, *19*, R90–R97.

(2) Di Prospero, N. A.; Fischbeck, K. H. *Nat. Rev. Genet.* **2005**, *6*, 756–767.

(3) Liquori, C. L.; Ricker, K.; Moseley, M. L.; Jacobsen, J. F.; Kress, W.; Naylor, S. L.; Day, J. W.; Ranum, L. P. *Science* **2001**, *293*, 864–867.

(4) Gatchel, J. R.; Zoghbi, H. Y. *Nat. Rev. Genet.* **2005**, *6*, 743–755.

(5) Schoser, B.; Timchenko, L. *Current genomics* **2010**, *11*, 77.

(6) Pearson, C. E. *PLoS Genet.* **2011**, *7*, e1002018.

(7) Zu, T.; Gibbens, B.; Doty, N. S.; Gomes-Pereira, M.; Huguet, A.; Stone, M. D.; Margolis, J.; Peterson, M.; Markowski, T. W.; Ingram, M. A. C.; Nan, Z.; Forster, C.; Low, W. C.; Schoser, B.; Somia, N. V.; Clark, H. B.; Schmechel, S.; Bitterman, P. B.; Gourdon, G.; Swanson, M. S.; Moseley, M.; Ranum, L. P. W. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 260–265.

(8) Gao, Z.; Cooper, T. A. *Hum. Gene Ther.* **2013**, *24*, 499–507.

(9) Wheeler, T. M.; Sobczak, K.; Lueck, J. D.; Osborne, R. J.; Lin, X.; Dirksen, R. T.; Thornton, C. A. *Science* **2009**, *325*, 336–339.

(10) Kalsotra, A.; Singh, R. K.; Gurha, P.; Ward, A. J.; Creighton, C. J.; Cooper, T. A. *Cell Reports* **2014**, *6*, 336–345.

(11) Pearson, C. E. *PLoS Genet.* **2011**, *7*, e1002018.

(12) McMurray, C. T. *Nat. Rev. Genet.* **2010**, *11*, 786–799.

(13) de Haro, M.; Al-Ramahi, I.; De Gouyon, B.; Ukani, L.; Rosa, A.; Faustino, N. A.; Ashizawa, T.; Cooper, T. A.; Botas, J. *Hum. Mol. Genet.* **2006**, *15*, 2138–2145.

(14) Kanadia, R. N.; Shin, J.; Yuan, Y.; Beattie, S. G.; Wheeler, T. M.; Thornton, C. A.; Swanson, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11748–11753.

(15) Chamberlain, C. M.; Ranum, L. P. W. *Hum. Mol. Genet.* **2012**, *21*, 4645–4654.

(16) Iyer, R. R.; Pluciennik, A.; Napierala, M.; Wells, R. D. *Annu. Rev. Biochem.* **2015**.

(17) Liu, G.; Leffak, M. *Cell & Bioscience* **2012**, *2*, 7.

(18) Castel, A. L.; Cleary, J. D.; Pearson, C. E. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 1–6.

(19) Gomes-Pereira, M.; Monckton, D. G. *Mutat. Res. - Fund. Mol. M.* **2006**, *598*, 15–34.

(20) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. *Nucl. Acids Res.* **2009**, *37*, 4149–4156.

(21) Childs-Disney, J. L.; Yildirim, I.; Park, H.; Lohman, J. R.; Guan, L.; Tran, T.; Sarkar, P.; Schatz, G. C.; Disney, M. D. *ACS Chem. Biol.* **2014**, *9*, 538–550.

(22) Warf, M. B.; Berglund, J. A. *RNA* **2007**, *13*, 2238–2251.

(23) Haghighat Jahromi, A.; Honda, M.; Zimmerman, S. C.; Spies, M. *Nucl. Acids Res.* **2013**, *41*, 6687–6697.

(24) Laurent, F. X.; Sureau, A.; Klein, A. F.; Trouslard, F.; Gasnier, E.; Furling, D.; Marie, J. *Nucl. Acids Res.* **2012**, *40*, 3159–3171.

(25) Pettersson, O. J.; Aagaard, L.; Andrejeva, D.; Thomsen, R.; Jensen, T. G.; Damgaard, C. K. *Nucl. Acids Res.* **2014**, *42*, 7186–7200.

(26) Wheeler, T. M.; Leger, A. J.; Pandey, S. K.; MacLeod, A. R.; Nakamori, M.; Cheng, S. H.; Wentworth, B. M.; Bennett, C. F.; Thornton, C. A. *Nature* **2012**, *488*, 111–115.

(27) Lee, J. E.; Bennett, C. F.; Cooper, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4221–4226.

(28) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller,

B. L. *J. Am. Chem. Soc.* **2008**, *130*, 16254–16261.

(29) Warf, M. B.; Nakamori, M.; Matthys, C. M.; Thornton, C. A.; Berglund, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18551–18556.

(30) Siboni, R. B.; Bodner, M. J.; Khalifa, M. M.; Docter, A. G.; Choi, J. Y.; Nakamori, M.; Haley, M. M.; Berglund, J. A. *J. Med. Chem.* **2015**, *58*, 5770–5780.

(31) Siboni, R. B.; Nakamori, M.; Wagner, S. D.; Struck, A. J.; Coonrod, L. A.; Harriott, S. A.; Cass, D. M.; Tanner, M. K.; Berglund, J. A. *Cell Reports* **2015**, *13*, 2386–2394.

(32) Pushechnikov, A.; Lee, M. M.; Childs-Disney, J. L.; Sobczak, K.; French, J. M.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2009**, *131*, 9767–9779.

(33) Lee, M. M.; Pushechnikov, A.; Disney, M. D. *ACS Chem. Biol.* **2009**, *4*, 345–355.

(34) Parkesh, R.; Childs-Disney, J. L.; Nakamori, M.; Kumar, A.; Wang, E.; Wang, T.; Hoskins, J.; Tran, T.; Housman, D.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 4731–4742.

(35) Rzuczek, S. G.; Gao, Y.; Tang, Z.-Z.; Thornton, C. A.; Kodadek, T.; Disney, M. D. *ACS Chem. Biol.* **2013**, *8*, 2312–2321.

(36) Childs-Disney, J. L.; Parkesh, R.; Nakamori, M.; Thornton, C. A.; Disney, M. D. *ACS Chem. Biol.* **2012**, *7*, 1984–1993.

(37) Guan, L.; Disney, M. D. *Angew. Chem. Int. Ed.* **2013**, *52*, 10010–10013.

(38) Guan, L.; Disney, M. D. *Angew. Chem. Int. Ed.* **2012**, *52*, 1462–1465.

(39) Rzuczek, S. G.; Park, H.; Disney, M. D. *Angew. Chem. Int. Ed.* **2014**, *53*, 10956–10959.

Chapter 2

Biological activities of CUG^{exp}-targeting ligands[1]

2.1. Introduction

One approach for DM therapeutic development is to target the deleterious interactions of MBNL1 and CUG^{exp} or $CCUG^{exp}$ using small molecules. To disrupt the MBNL1-CUG^{exp} complex, small molecules can target either MBNL1 or toxic RNA transcripts in which CUG^{exp} and $CCUG^{exp}$ were believed to be preferred targets because their locations in non-coding regions. A screening assay done by the Disney group revealed that small molecule binding to MBNL1 caused DM1 disease phenotype, which supports the strategy of targeting CUG^{exp}.[1] MBNL1-CUG^{exp}/ $CCUG^{exp}$ small-molecule inhibitors developed by other groups are based on hits from library screening and are discussed in Chapter 1. Our group is interested in developing small molecules that disrupt MBNL1-CUG^{exp}/ $CCUG^{exp}$ using a rational design approach based on reported crystal structures of the toxic RNA transcript. In addition, we are attracted to this strategy because of a lack of small molecules with the capacity to target RNA with potency and selectivity. The A-form hairpin of CUG^{exp} containing repetitive UU mismatches separated by two GC base pairs[2-4] is distinct from other cellular single-strand RNA transcripts. Thus, finding agents to selectively bind CUG^{exp} could provide an important answer to the question of whether RNA is a druggable target. To design CUG^{exp}-targeting small molecules, we have

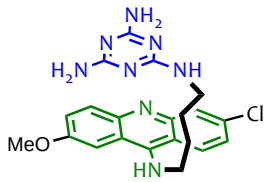
---

[1] Some of the material in this chapter was adapted from the following publications:
Jahromi, A. H.; **Nguyen, L**.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. *ACS Chem. Biol.* **2013**, *8*, 1037–1043.
Jahromi, A. H.; Fu, Y.; Miller, K. A.; **Nguyen, L.;** Luu M. L.; Baranger, A. M.; Zimmerman, S. C. *J. Med. Chem.* **2013**, *56*, 9471–9481.
Wong, C.-H.; **Nguyen, L**.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Ho, T.; Chan, E. H. Y.; Chan, W.-Y.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 6355−6361.

conjugated a UU mismatch recognition unit and an RNA binding unit. This strategy is not only simple but has also provided several relatively potent and selective small molecules for targeting CUG^{exp}, and the same general approach has been applied successfully to DM2.

2.2. Development of intercalating ligands targeting CUG^{exp}

In 2009, our laboratory reported a simple molecule targeting CUG^{exp} that was rationally designed based on the reported X-ray structure of $(CUG)_{12}$.[5] Ligand **1** (Figure 2.1) was designed as a conjugate of a UU Janus-Wedge-recognition melamine motif and a known DNA and RNA-binding acridine unit that provides a hydrophobic driving force for



1

Figure 2.1. The first triaminotriazine-acridine conjugate designed by our group for targeting DM1

binding. The triaminotriazine unit was proposed to form a base triplet with a UU mismatch or to induce a base flipping of a uracil while pairing the other uracil (Figure

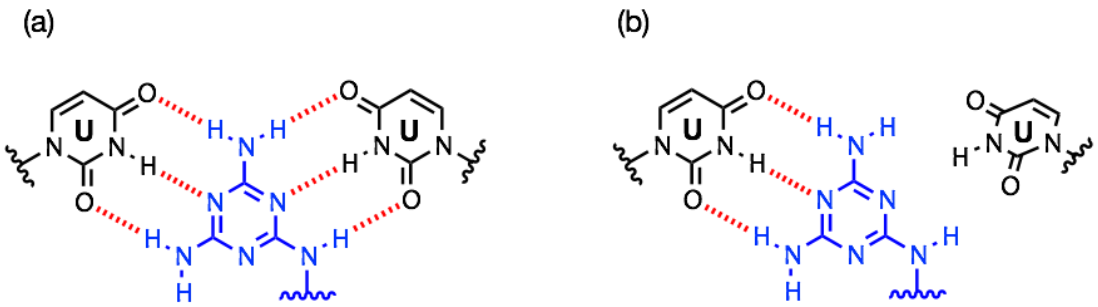
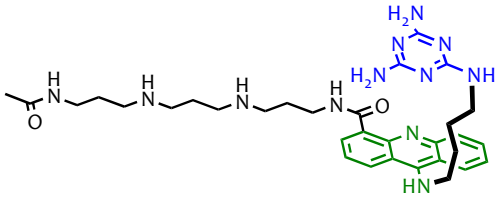


Figure 2.2. Proposed binding mode of a triaminotriazine unit and a UU mismatch. (a) Base triplet model. (b) Uracil flipping model.

2.2). This ligand was found to bind to *d*(CTG)$_2$ and *r*(CUG)$_2$ with low micromolar binding affinities (K_D) as measured by isothermal titration calorimetry (ITC) and to inhibit MBNL1-(CUG)$_{12}$ with an IC_{50} of 46 μM determined using an electrophoresis mobility shift assays (EMSA).[5]

However, cell-based experiments showed that **1** is not only poorly cell permeable but also highly toxic, possibly the result of non-specific intercalation of the acridine unit in the unstacked form with other DNA and RNA sequences. To improve the cell uptake, Dr. Haghighat Jahromi synthesized water-soluble derivative **2** (Figure 2.3) by conjugating a polyamine to **1**.[6] The presence of the polyamine tail can help **2** be easily transported through cellular membranes via the polyamine transporting system (PTS). The biological activity of **2** was studied *in vitro*, showing that in comparison to **1**, ligand **2** maintains the binding affinity and inhibition potency toward the CUG^{exp}-MBNL1 complex. In addition, the cytotoxicity of 2 was reduced in comparison to **1**.



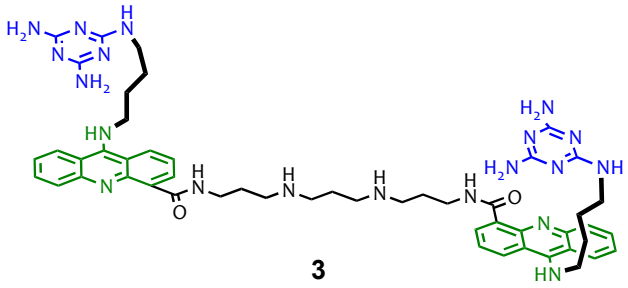
2

Figure 2.3. Ligand **2**, a derivative of **1** with a polyamine tail for improving water solubility and cellular uptake

To enhance the inhibition potency of ligands **1** and **2**, Dr. Haghighat Jahromi developed dimeric ligands targeting CUG^{exp} by linking two core units of **1** by polyether or polyamine linkers. The length of linkers was varied to optimize the bivalent effect and find the best dimeric ligand. Of ten dimers prepared, **3** (Figure 2.4) was the most active in a surface plasmon resonance (SPR) assay with 206-fold stronger binding affinity and

266-fold greater inhibition potency toward CUG^{exp} and CUG^{exp}-MBNL1 complexes, respectively, in comparison with monomer **2**.

The in-cell biological activities of ligands **2** and **3,** including their cytotoxicity and the ability of disrupting DM1 disease foci and rescuing the splicing defect will be described in Section 2.3.



3

Figure 2.4. The most potent dimeric ligand.

2.3. Cellular bioactivities of ligands 2 and 3

2.3.1. Toxicity study of 2 and 3 using Sulforhodamine B colorimetric assay

Sulforhodamine B colorimetric assays were performed to study how ligands affect cell viability.[7] Sulforhodamine B, a bright-pink aminoxanthene dye, is known to bind to proteins. After treating with ligands, fixed cells were incubated with Sulforhodamine B. The amount of Sulforhodamine B bound to cellular proteins depends on the protein

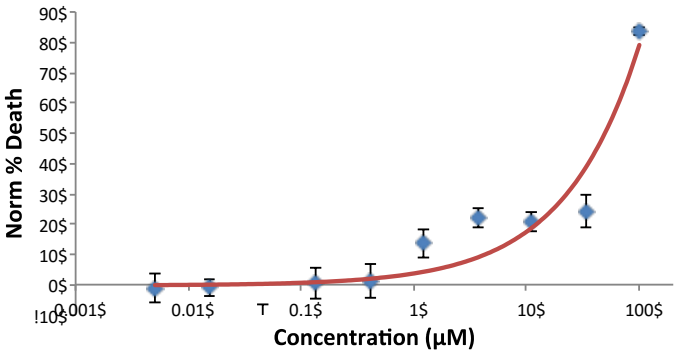


Figure 2.5. Toxicity curve for ligand **2**. Semi-logarithmic plot of percent HeLa cell death against log [**2**]. Error bars represent standard deviations from three independent experiments. Something about conditions should be said (pH, buffer and ionic strength, etc) or indicate that these can be found in some other section.

content of the cells, so this assay provides information about cell growth. HeLa cells were incubated with serial dilutions of **2**. Cells that survived after 24-h incubation were stained with Sulforhodamine B. Fluorescence of bound Sulforhodamine B was measured and the percentage of cell death under ligand treatment was calculated from a comparison with untreated samples. Plotting the percentage of cell death against tested concentrations of **2** provides a toxicity curve with an IC_{50} of 50 μM, which is the concentration resulting in 50% cell death (Figure 2.5). A similar experiment was performed with **3**, showing less than 20% cell death at 75 μM after 24-h incubation. However, longer incubations (36 h, 48 h) showed that **3** was significantly more toxic than **2** with an IC_{50} of less than 10 μM.

2.3.2. Foci disruption by ligands 2 and 3 using fluorescence *in situ* hybridization

The hallmark of DM1 disease is a sequestration of MBNL proteins by CUG repeats to form disease foci in the nuclei of DM1 cells. If the ligand binds to $(CUG)_n$ and inhibits the formation of the MBNL1-$(CUG)_n$ complex, the ribonuclear foci should disappear or not form in the first place allowing MBNL1 protein to disperse throughout the nucleus. To determine if ligands **2** and **3** are able to dissolve the DM1 foci, fluorescence *in situ* hybridization (FISH) experiments were performed. HeLa cells were co-transfected with DT960 plasmid containing truncated DMPK-$(CTG)_{960}$ and GFP-MBNL1 plasmid co-expressing GFP and MBNL1 proteins that are covalently linked to generate a DM1 model cells. Thus, the fluorescence of GFP was utilized to locate MBNL1, while the FISH probe, 5'-Cy3-$(CAG)_{10}$, was used to stain CUG^{exp}. Transfected HeLa cells were treated with ligands **2** and **3**, and spermine (*N*-[3-({3-[(3-aminopropyl)amino]propyl}-amino)propyl] acetamide) as a negative control. The number of foci in each sample were counted, and then compared to untreated DT0 and DT960 controls.

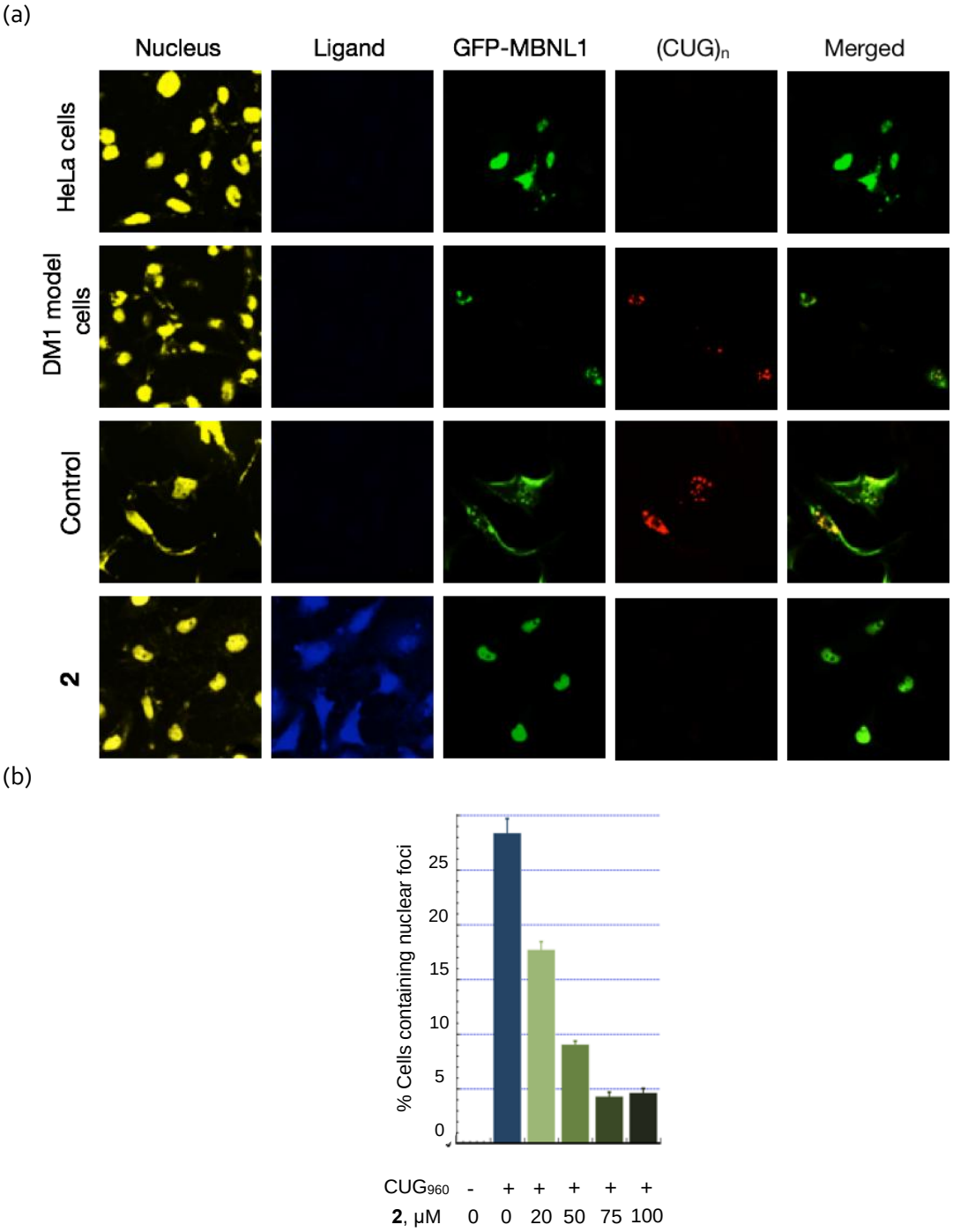


Figure 2.6. Ligand **2** dissolves the disease foci in DM1 model cells in FISH experiments. (a) The disease foci present in DM1 model cells (row 2) as well as in a negative control sample (50 μM) (row 3). Ligand **2** disrupts the disease foci at two representative concentrations, 50 and 75 μM after 48 h incubation (rows 4 and 5). Each box is (150x150) μm. (b) Plot of the number of $(CUG)_n$ foci-containing cells versus different ligand concentrations by counting over 100 cells. The error bars are standard deviations of at least three independent experiments.

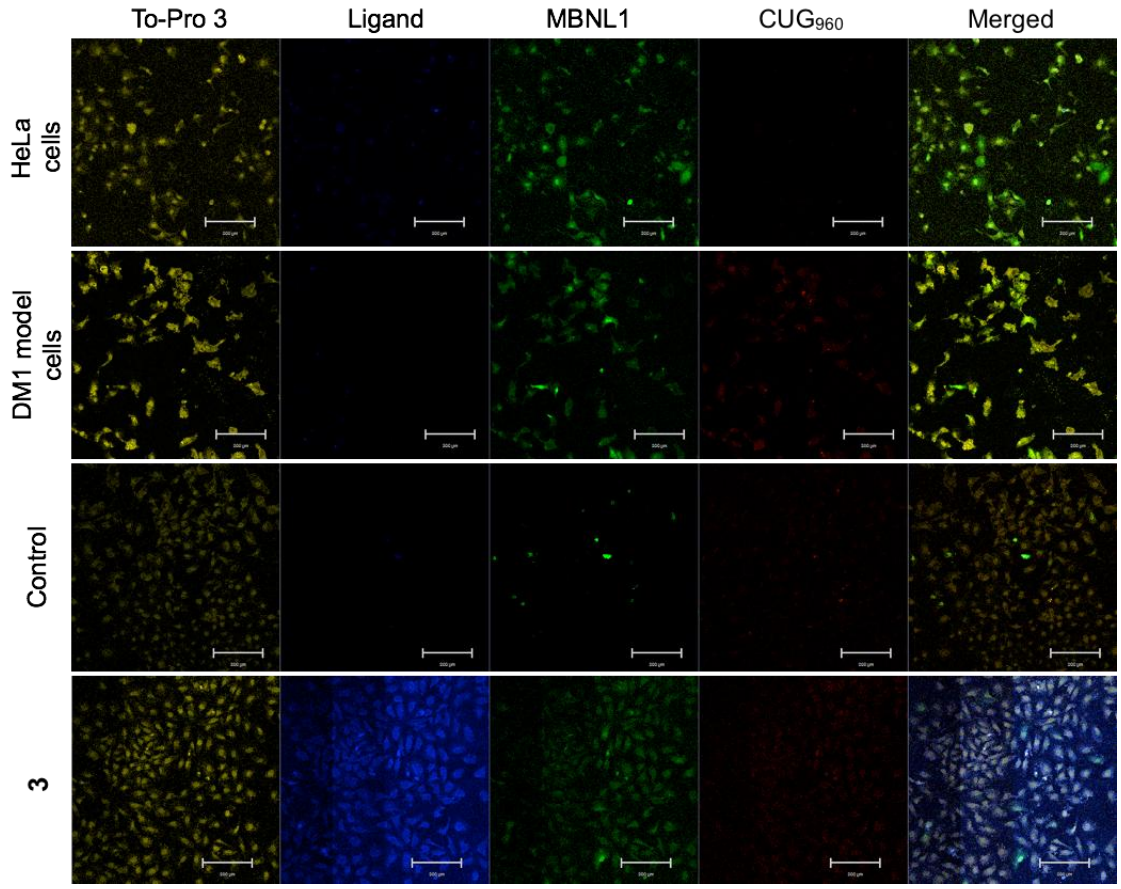


Figure 2.7. Foci dispersion in DM1 cells treated with dimeric ligand **3** (50 μM, 36 h incubation). To-Pro 3, compound, GFP-MBNL1, and Cy3-CAG_{10} channels were observed by excitation at 633 nm, 405 nm, 488 nm and 564 nm, respectively.

For cells transfected with CUG_0 and GFP-MBNL1 plasmids, no Cy3 signal of FISH probe was observed because 5'-Cy3-$(CAG)_{10}$ had no targets and was easily washed out in the washing steps resulting in much lower intensities compared to the signal observed for the CUG_{960} control. In addition, the GFP-MBNL1 signal was fully dispersed in the whole nucleus because there was no CUG_{960} to sequester GFP-MBNL1 (Figure 2.6 and Figure 2.7). In contrast, GFP-MBNL1 sequestration by CUG^{exp} in the DM1 model cells was confirmed with a co-localization of Cy3-CAG_{10} (red signal) and GFP (green signal) as the yellow dots in nuclei (merged column, Figure 2.6 and Figure 2.7).

Taking advantage of the inherent fluorescence of the acridine moiety, the ligands were observed to spread throughout the whole cell (Figure 2.6 and Figure 2.7), providing

direct evidence of cell permeability. Approximately 75% of DM1 foci were dispersed for cells incubated with **2** for 48 h treatment (Figure 2.6b), whereas ligand **3** showed better cellular efficacy with 100% foci disruption after 36-h incubation (Figure 2.7). In contrast, no foci dispersion was observed for the cells treated with spermine control, supporting the important role of the recognition and binding units in the activity of ligands **2** and **3**.

2.3.3. Tracking GFP-MBNL1 foci dispersion by 2 and 3 using live cell imaging

The promising results from the FISH study encouraged us to pursue live cell imaging experiments that utilized the fluorescence of acridine and GFP. Compared to the FISH experiment that provided information only after fixing the cells, live-cell imaging is able to look at single live cells and track their cellular events over time. Consequently, this approach can provide direct evidence of the GFP-MBNL1 dispersion from identified foci simply by monitoring the fluorescent signal of GFP.

At the time ligands **2** and **3** were added, no signal was observed under the acridine channel, whereas bright green dots were observed using the GFP channel (488 nm) showing the sequestration of GFP-MBNL1 in the nucleus. After 2–7 h treatment, the acridine signal was obseved with excitation at 405 nm, which showed the cellular pernetration of the compounds, whereas the green signal of GFP-MBNL1 in puntate foci observed at t = 0 h gradually disappeared to be replaced by green GFP signal covering the whole nucleus. This observation provides direct evidence that the ligands penetrated both the cell and nuclear membranes and, outcompeted the GFP-MBNL1 for CUG^{exp} binding. This inhibition resulted in GFP-MBNL1 liberation and its spreading throughout the nuclei (Figure 2.8a and Figure 2.9).

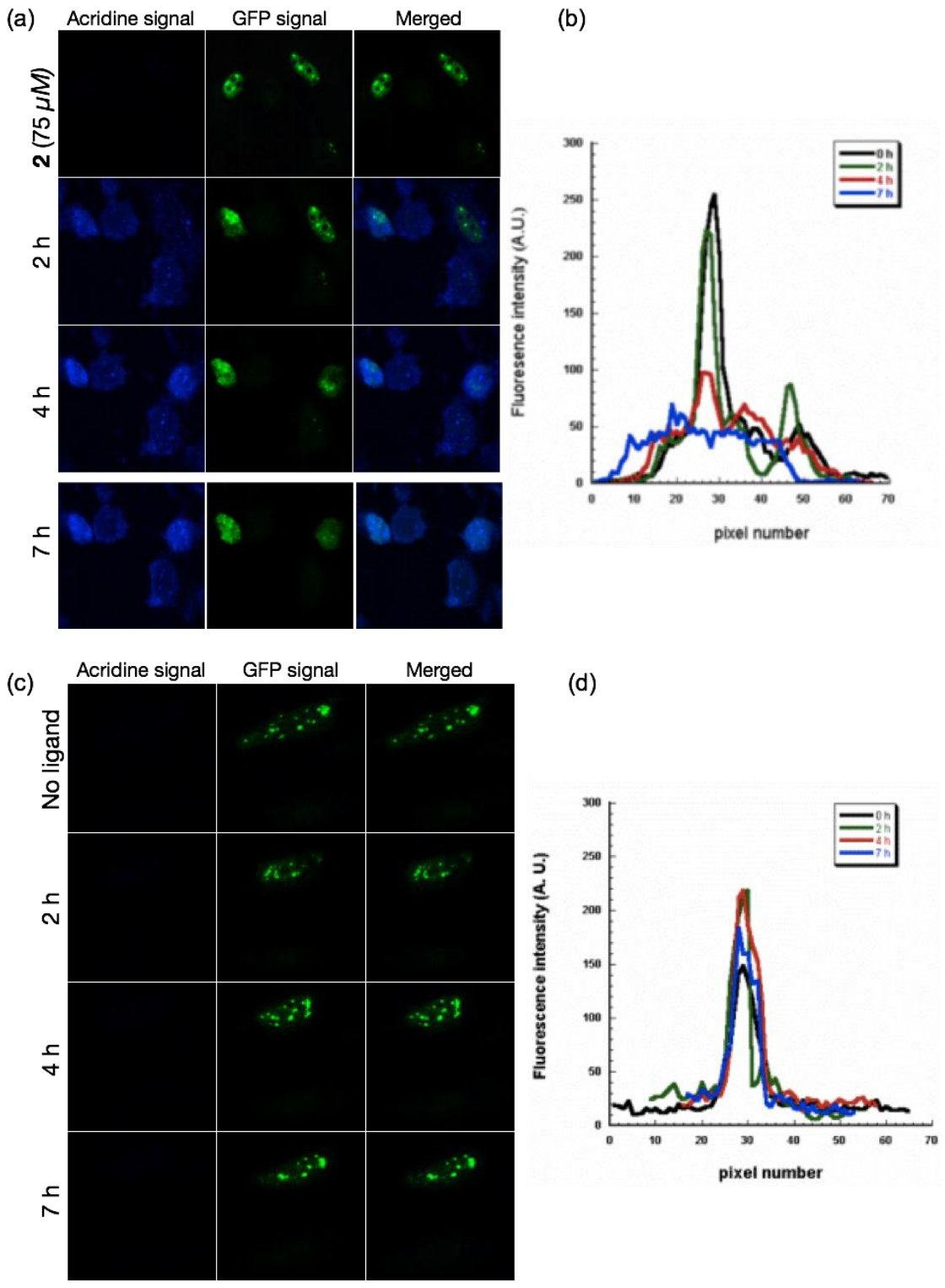


Figure 2.8. Tracking GFP-MBNL1 dispersion under the treatment of **2**. *DM1 model cells treated with **2** (75 μM)*: (a) GFP-MBNL1 dispersion versus time. (b) A decrease of GFP fluorescence intensity by time. *DM1 model cells without treatment of ligand*: (c) No GFP-MBNL1 dispersion observed. (d) GFP fluorescence intensity was unchanged.

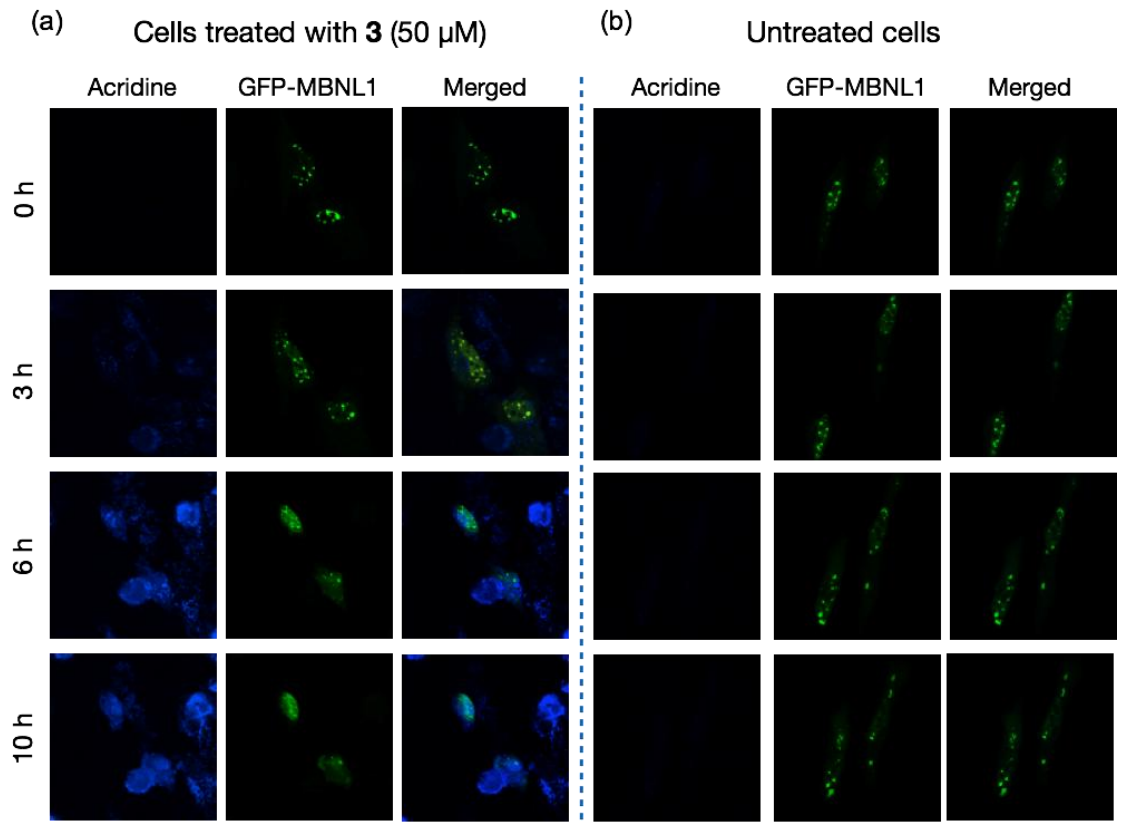


Figure 2.9. Tracking GFP-MBNL1 dispersion under the treatment of **3**: (a) GFP-MBNL1 dispersion versus time in DM1 model cells treated with **3** (50 μM). (b) No GFP-MBNL1 dispersion observed in DM1 model cells without treatment of ligand.

In constrast, in a control experiment without compound, dispersion of GFP signal was not observed over time and there was no signal observed when cells were excited at 405 nm (Figure 2.8b). To quantify the change in GFP signal at the GFP-MBNL1 foci at least qualitatively, we plotted the change in fluorescence at representative foci. In ligand-treated samples, there was a decrease in GFP fluorescence by time from 250 (t = 0 h) to 50 (t = 7 h) fluorescence units (Figure 2.8c), wheras in untreated samples that signal remained mostly unchanged (Figure 2.8d).

2.3.4. Reversal of *IR* pre-mRNA splicing defect in DM1 model cells by ligand 2

The DM1 phenotype is related to splicing defects of a set of pre-mRNAs[8] including insulin receptor (*IR*) resulting from a decrease in cellular levels of MBNL1.[8,9] The *IR*

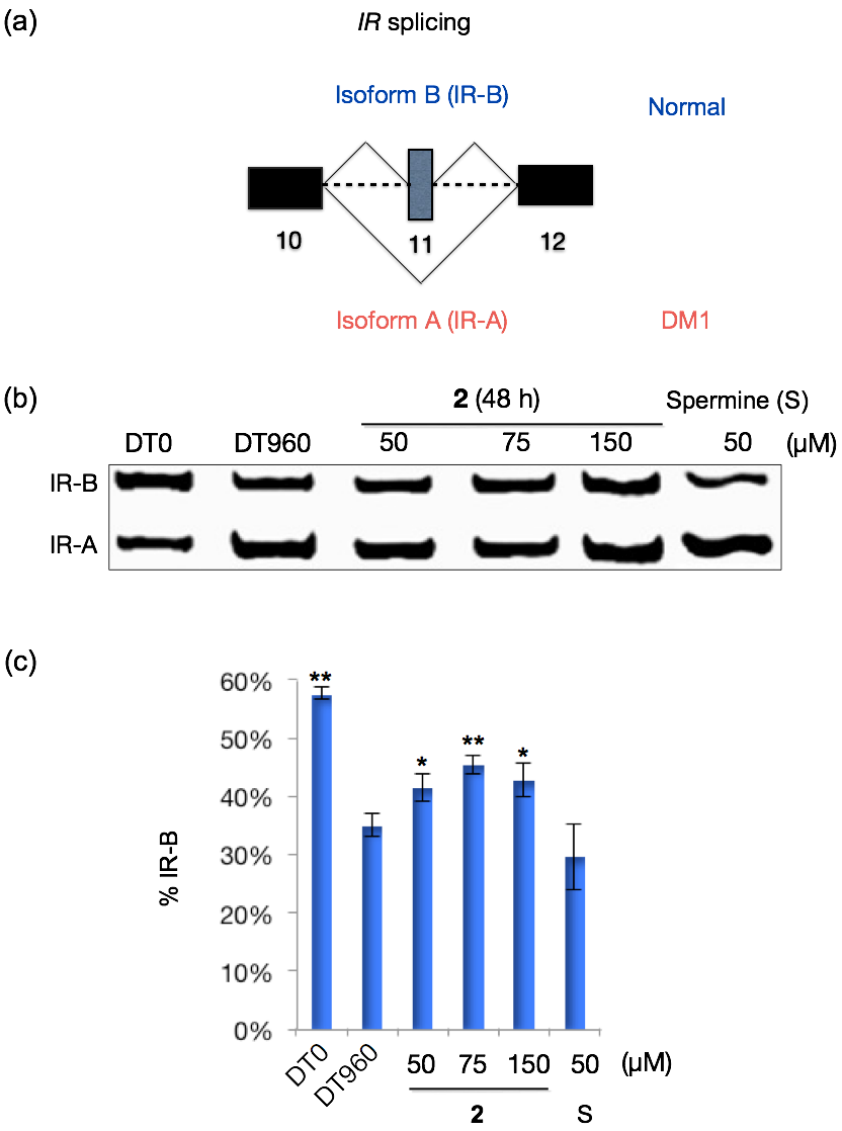


Figure 2.10. Reversal of *IR* pre-mRNA splicing defects by ligand **2**. (a) *IR* pre-mRNA missplicing in DM1. (b) *IR* splicing gel showing **2** rescued *IR* splicing defect at 50, 75, and 150 μM. (c) Percent of isoform B in splicing experiments. Error bars represent standard deviations of at least three independent experiments. * $P < 0.05$; ** $P < 0.005$ (two tailed *t*-test).

splicing defect in DM1 is the dominance of the exon 11 skipped isoform A in comparison to the exon 11 included isoform B (Figure 2.10a). Because ligands **2** and **3** disrupted the interaction of MBNL1 and CUG^{exp} in DM1 model cells, we sought to perform splicing assays to examine if ligands can reverse the splicing defect of the *IR* minigene. Thus, **2**

and **3** were treated with DM1 model cells as HeLa cells co-transfected with DT960 plasmid and *IR* minigene-containing plasmid.

Ligand **2** rescued about 40% of the splicing defects of the *IR* minigene pre-mRNA (Figure 2.10b and 2.10c), whereas no correction in *IR* splicing defects was observed for a spermine control. This result was consistent with confocal experiments in which **2** disrupted MBNL1-CUG^{exp} foci but spermine did not. There was no statistically significant difference in *IR* splicing-defect correction with DM1 model cells treated with **2** at different concentrations (50, 75 and 150 μM), but 75 μM was found as the best concentration taking into account both the *IR* splicing-defect rescue and the compound cytotoxicity (Figure 2.10b and c).

In splicing experiments, ligand **3** was highly toxic after 20-24 h incubation with DM1 model cells with dots inside cells observed under normal microscopy, not allowing further splicing assays with 48 h incubation. Consequently, the protocol was edited to decrease the incubation time to 20–24 h, but the following step in the assay could not be performed because of the poor quality of the isolated total RNA. The RNA was considered as being in good condition with an A_{260}/A_{280} of about 2.0 and was acceptable if that ratio was above 1.6, but the isolated RNA from cells treated with ligand **3** was around 1.2. Several samples of RT-PCR trials were performed with low-quality isolated RNA, but the bands on the splicing-experiment gel were faint and the results were inconsistent.

2.4. Development of groove-binding ligands targeting CUG^{exp}

The main goal of the project is to develop a new set of CUG^{exp}-targeting ligands with comparable CUG^{exp}-binding affinity and MBNL1-CUG^{exp}-inhibition potency and an

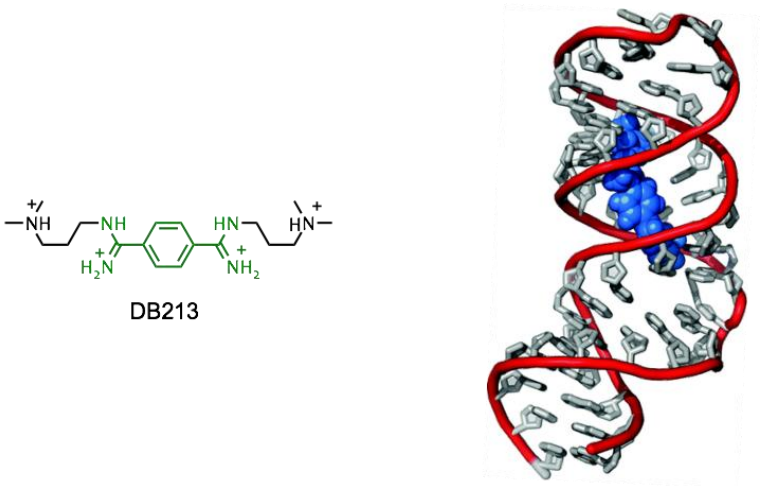


Figure 2.11. Structure of DB213 and the NMR structure of the compound on HIV-1 FS RNA reported by Butcher.[11] The bisamidinium unit is highlighted in green.

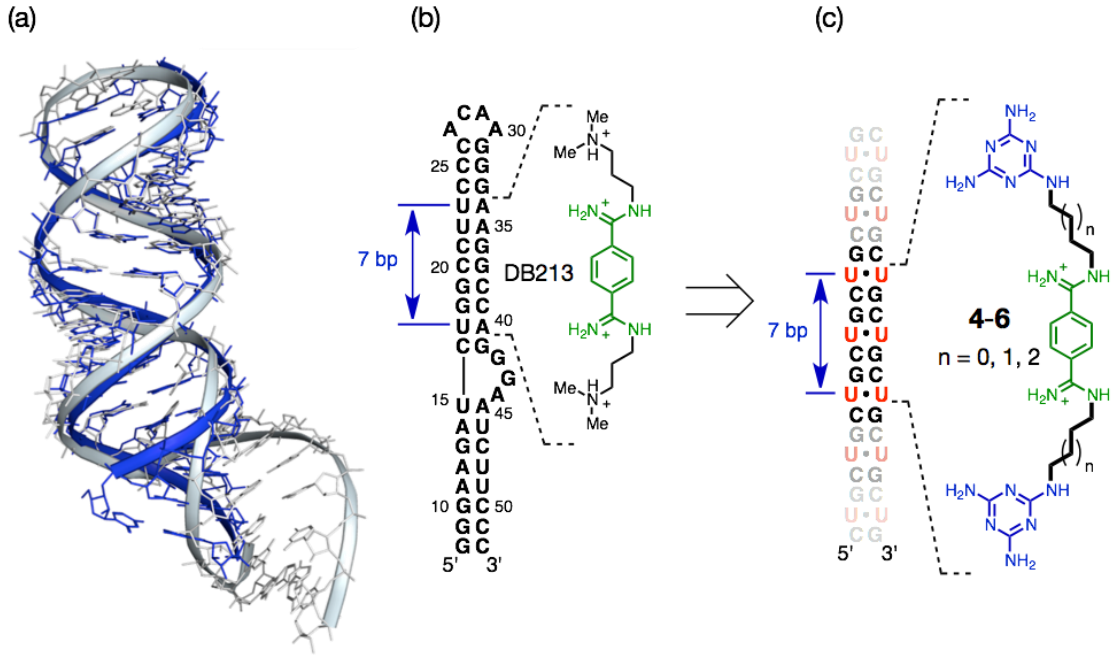
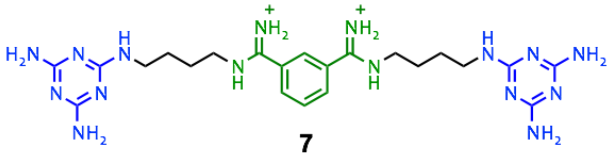


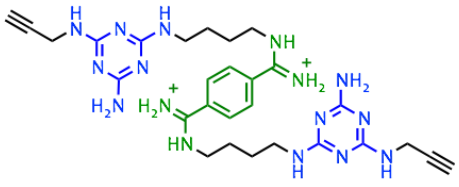
Figure 2.12. Design of bisamidinium-based ligands targeting CUG^{exp}. (a) Overlaying structures of HIV fs RNA and CUG^{exp} (PDB: 3GM7). (b) DB213 covers seven base pairs on the HIV-1 FS RNA. (c) CUG^{exp}-targeting bisamidinium-based ligands were designed by replacing dimethylammonium groups with triaminotriazine units.

improvement in cytotoxicity, water-solubility, and cellular uptake in comparison with the acridine-based ligands discussed above. Thus, Dr. Wong in our group designed another

CUG^{exp} binder based on a known RNA groove binder, DB213 (Figure 2.11).[10] In 2011, Butcher and co-workers reported the NMR structure of DB213 on the major groove of HIV-1 frameshift site RNA (HIV fs RNA)[11] that adopts a A-form hairpin structure. Overlaying the NMR structure of HIV fs RNA[11] and the crystal structure of $(CUG)_6$ (PDB: 3GM7) shows many structural similarities (Figure 2.12a). Interestingly, DB213 covers seven base pairs on HIV fs RNA that is exactly the distance between every third UU mismatch (Figure 2.12b). Thus, Dr. Wong designed a set of bisamidinium-based ligands in which the two dimethylamino groups of DB213 were replaced by triaminotriazine units for targeting UU mismatches. The aliphatic linker between the triaminotriazine and bisamidinium units was varied from three to five carbon atoms to optimize the activity of the ligand (Figure 2.12c). Isothermal titration calorimetry (ITC) studies done by Dr. Wong showed that ligand **5** exhibited the strongest binding affinity toward $(CUG)_{12}$ with K_D = 8 μM. To examine if substituent positions on the benzene ring would affect the binding affinity and activities of ligand, ligand **7** (Figure 2.13), a meta-substituted analog of **5**, was synthesized. Ligand **7** showed a similar binding affinity



7



8

Figure 2.13. Ligands **7** and **8**, a meta-substituted analog and an alkyne-containing derivative of **5**, respectively.

toward $(CUG)_{12}$ in comparison with **5**. It is important to note that all bisaminidum-based ligands are water soluble, thus organic solvent (e.g. DMSO) was not required in most of biological assays.

2.5. *In-vitro* biological activities of bisamidinium-based ligands

The inhibition potency of bisamidinium-based ligands toward the CUG^{exp}-MBNL1 complex was determined using electrophoresis mobility shift assays (EMSA). The first EMSA experiment was to measure the binding affinity of CUG^{exp} and a truncated

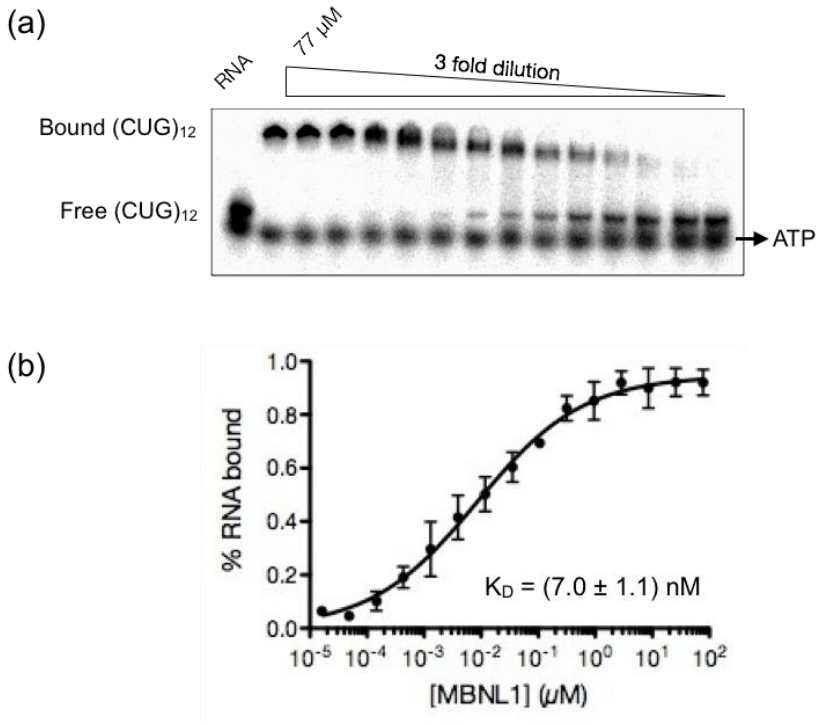


Figure 2.14. Binding study of MBNL1 and $(CUG)_{12}$. (a) EMSA binding gel. (b) Binding curve of MBNL1 and CUG^{exp}. Error bars represent standard deviations of three independent experiments. The detail of experiment can be found in section 2.8 (Materials and methods).

MBNL1. The ^{32}P-labeled $(CUG)_{12}$ was incubated with serial dilutions of MBNL1. The complexes and free RNA were separated on a PAGE gel (Figure 2.14a). Plotting the percentage of RNA bound against concentrations of MBNL1 gave a sigmoidal binding

curve (Figure 2.14b) that was fitted to a standard binding equation providing an K_D of 7 nM.

In the inhibition assays, each ligand serially diluted was incubated with MBNL1-$(CUG)_{12}$ complexes. Because of the very stringent conditions of the EMSA binding buffer (high salt buffer and the gel ran at 4 °C), 25% DMSO was required to obtain the highest concentration of ligands. The RNA fraction bound was normalized to the control which contained 25% DMSO. The normalized RNA fraction bound was plotted versus the log function of ligand concentrations showing that **5** was the most active ligand in EMSA experiments with IC_{50} = 115 μM, which is consistent with ITC binding results (Figure 2.15). The corresponding K_i of ligand **5** is 2.6 μM, which is comparable to that measured for the acridine-based ligands. EMSA runs under different conditions using a

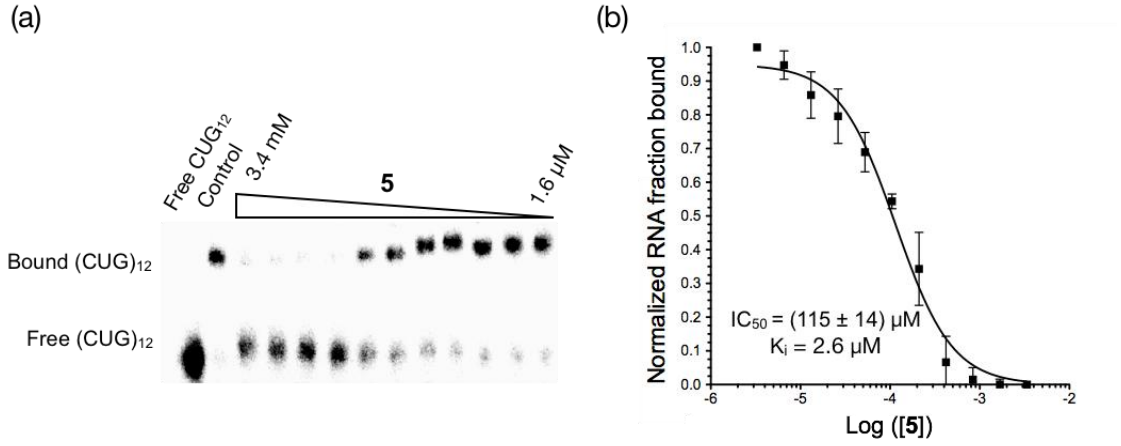


Figure 2.15. Inhibition study of ligand **5**. (a) Inhibition gel. (b) Inhibition curve of ligand **5**. Error bars represent standard deviations from three independent experiments. The detail of experiment can be found in section 2.8 (Materials and methods).

higher percentage of Triton-X added to eliminate aggregates of the ligand showed a similar K_i = 3.6 μM (done by Dr. Richardson). In addition, inhibition assays were performed with several reported MBNL1-CUG^{exp} inhibitors (e.g. pentamidine,[12,13] para-H1, and meta-H[14]), showing much less potency in comparison with ligand **5**.

The other bisamidinium based ligands **4** and **6** had weaker inhibition potencies (Figure 2.16). The *meta*-bisamidinium-melamine **7** showed a comparably strong inhibition to MBNL1-$(CUG)_{12}$ complex, but it is not as water soluble as **5**. Additionally, alkyne derivative **8** synthesized by Long Luu (Figure 2.13) has a similar IC_{50} in comparison to **5** (Figure 2.16), suggesting the alkyne (or propargyl) substituents do not affect the inhibition potency of ligand. The preliminary data for **8** revealed an important ligand for use in Fragment Based Drug Design (FBDD), *in situ* click chemistry on the $(CUG)_n$ template, and a range of other strategies.

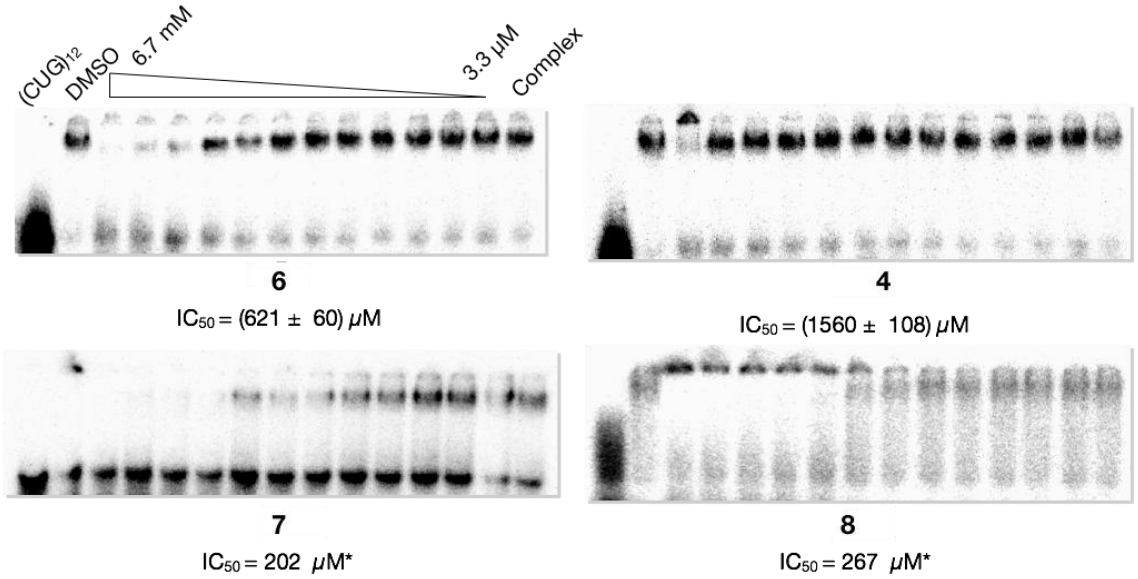


Figure 2.16. Inhibition study of bisamidinium-based ligands. IC50 values were obtained from three independent experiments, * from two independent experiments.

2.6. Cellular biological activities of bisamidinium-based ligands

2.6.1. Cytotoxicity study of ligands 4-8 using Sulforhodamine B colorimetric assay

Because toxicity is an important factor that dictates testing concentrations and, more fundamentally whether a ligand can be further developed at all, SRB experiments were carried out with DB213, **4**-**8**, and pentamidine following the protocol reported in the

section 2.3.1. The bisamidinium-based ligands showed significant toxicity in comparison to acridine-based ligands with IC_{50} > 100 μM for 3-d incubation (Figure 2.17). Additionally, toxicity studies were performed with mouse myoblasts and DM1 fibroblasts, showing a similar result with no toxicity. Importantly, further cytotoxicity

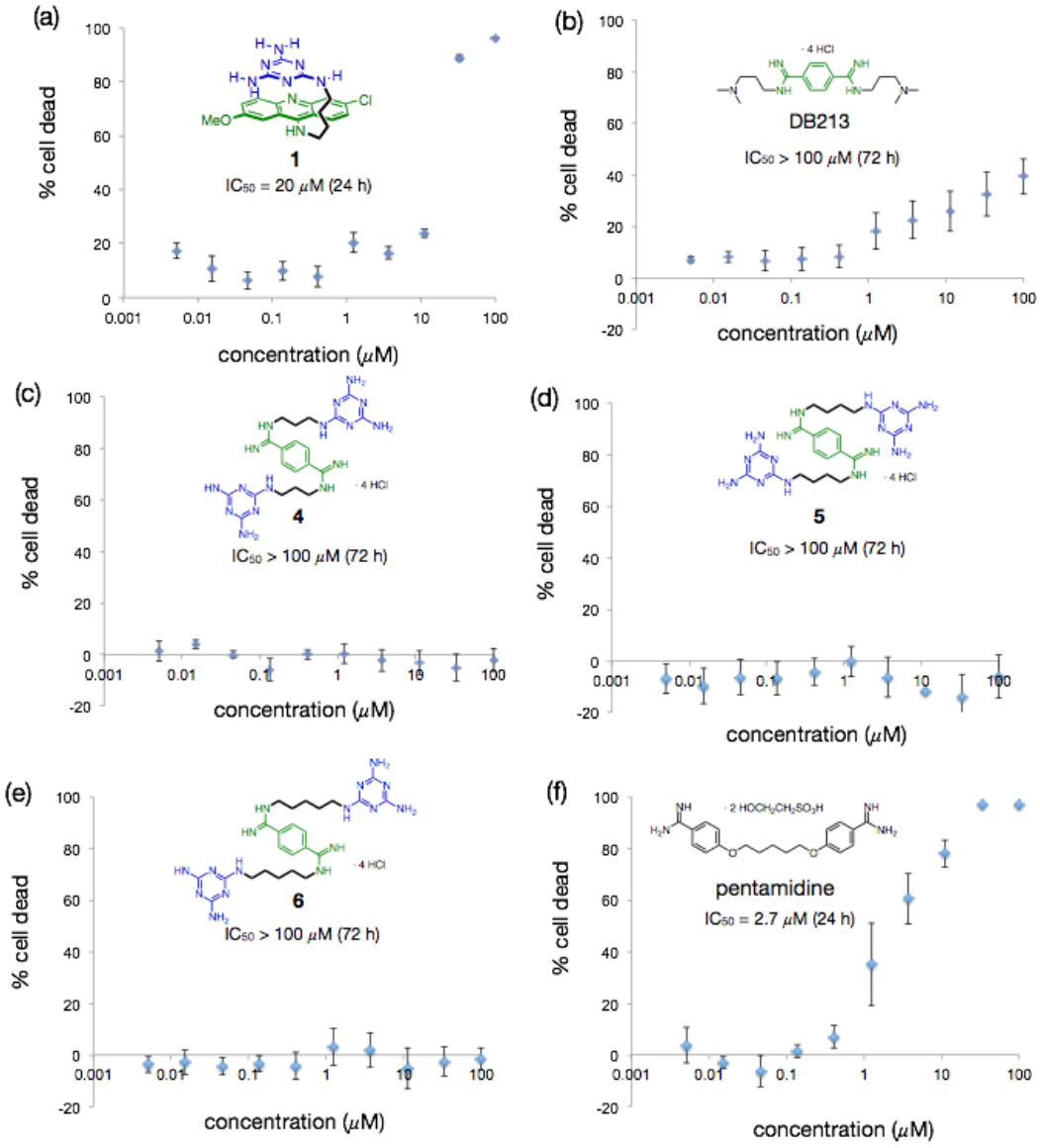


Figure 2.17. SRB toxicity profiles of (a) ligand **1** for 24 h, (b–e) DB213, ligands **4**–**6** for 72 h and (f) pentamidine for 24 h in HeLa cells. The detail of experiment can be found in section 2.8 (Materials and methods).

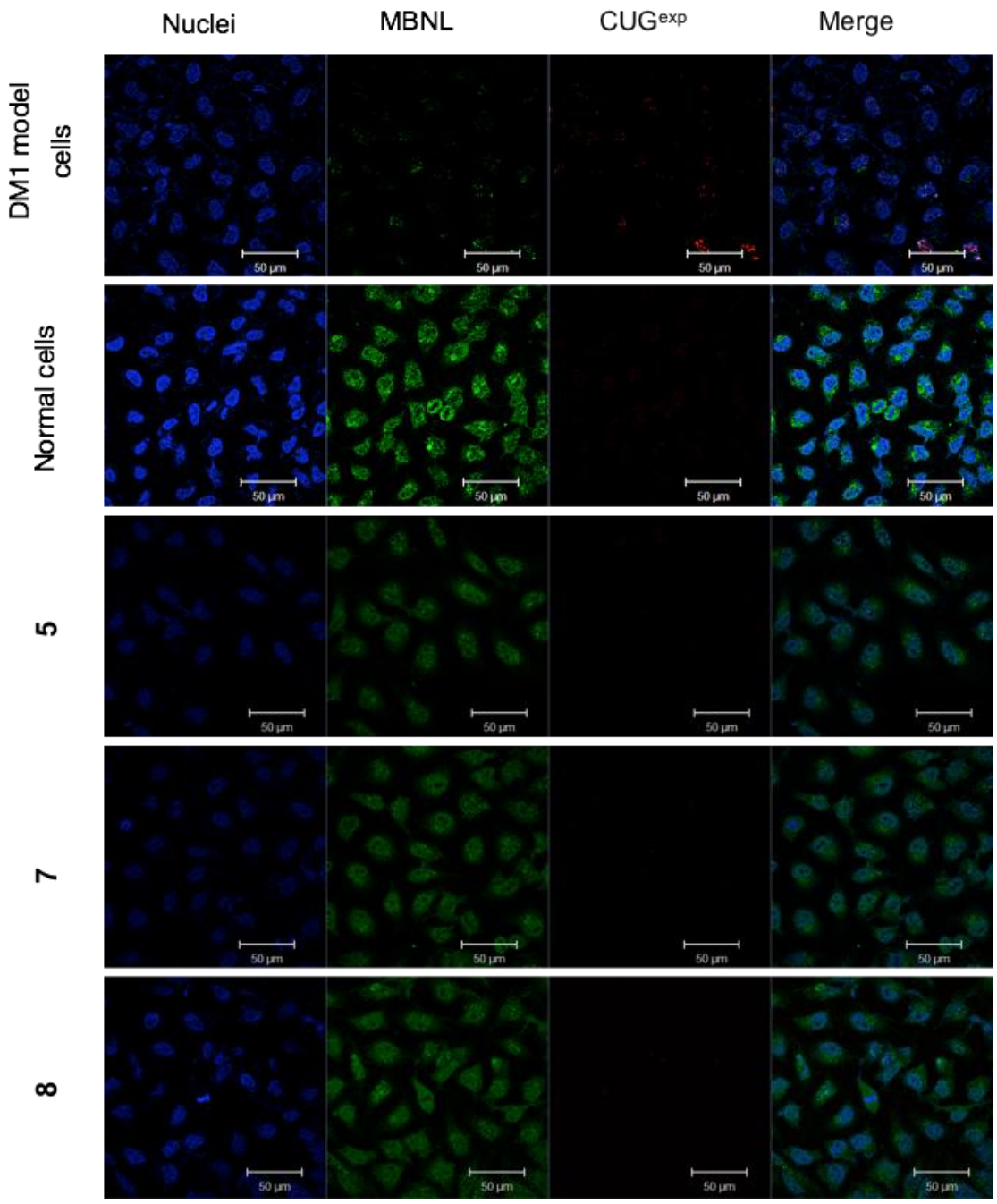


Figure 2.18. Foci dispersion by bisamidinium-based ligands (100 µM, 48 h). Very little detail here. Make sure you add the necessary detail or say explicitly that the detail is in the text or Experimental.

studies in mice revealed the MTD (maximum tolerated dose) of **5** in the range of 50 to 75 mg/kg (performed by Dr. Rachel Botham, Hergenrother group).

2.6.2. CUG^{exp}-MBNL1 foci disruption by bisamidinium-based ligands

The formation of MBNL1-CUG^{exp} foci is a key event in DM1. Thus, we performed FISH-immunofluorescence experiments to study if bisamidinium-based ligand can disrupt the interaction between MBNL1 and CUG^{exp} in DM1 model cells. HeLa cells were transfected with a DT960 or DT0 plasmid to generate DM1 model or normal cells, respectively. The CUG^{exp} was detected using a commercially available FISH probe, 5'-Cy3-$(CAG)_{10}$, whereas endogenous MBNL proteins were probed by anti-MBNL antibody followed with staining with an Alexa Fluro 488-containing secondary antibody.

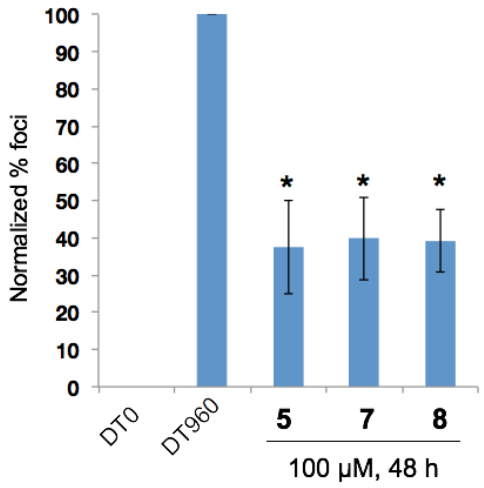


Figure 2.19. Quantification of ribonuclear foci area in DT0 or DT960-transfected HeLa cells treated with ligands **5**, **7**, and **8** at 100 μM for 48 h. Error bars represent standard deviations from three independent experiments. * $P < 0.05$ (two tailed *t*-test).

In untreated cells, punctate foci were clearly observed as a co-localization of CUG^{exp} and MBNL proteins in DM1 model cells, whereas there was no CUG^{exp} detected and MBNL proteins spread throughout the nucleus in normal cells (Figure 2.18, rows 1 and 2). DM1 model cells treated with **5**, **7**, and **8** at 100 μM showed a significant reduction in the number and size of disease foci for 48-h incubation (Figure 2.18, rows 3-5). In

particular, approximately 60% of foci dissipated (Figure 2.19). Further studies performed by Jessie Peh (Hergenrother group) showed that **5** disrupted the disease foci in a time and dose-dependent manner.[15]

2.6.3. Reversal of *IR* pre-mRNA splicing defects by ligand 5

To examine if **5** was able to rescue splicing defects, splicing assays for *IR* and *cTNT* pre-mRNAs were performed. For the *IR* splicing assay, HeLa cells were co-transfected with DT960 or DT0 and *IR* minigene plasmids were treated with **5** at 100 μM for 24, 48, and

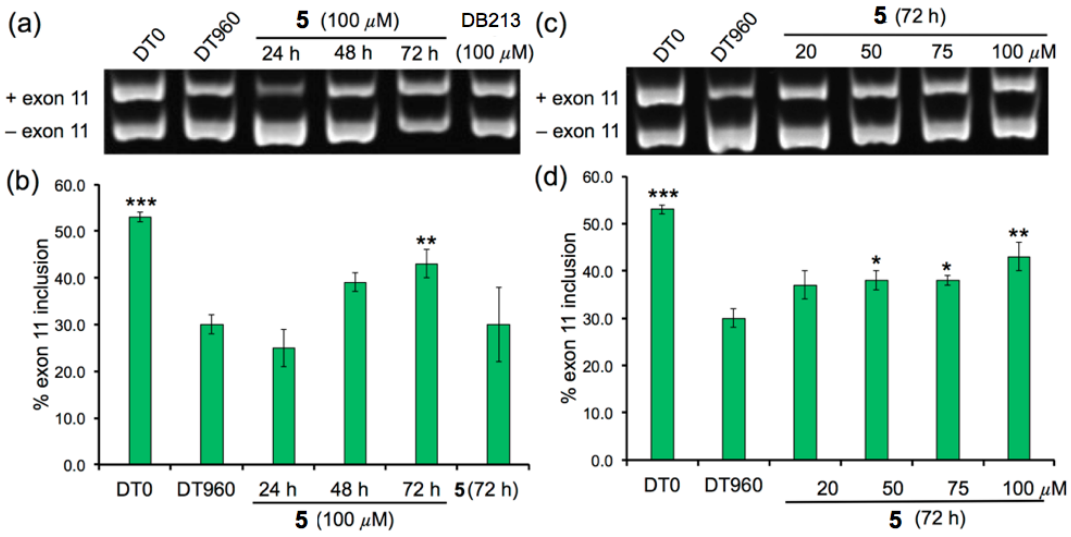


Figure 2.20. Reversal of IR pre-mRNA splicing defect by ligand **5**. (a) Gel of a time-dependent experiment. (b) Quantitative results of time-dependent experiments. (c) Gel of a dose-dependent experiment. (d) Quantitative results of dose-dependent experiments. Error bars represent standard errors of means from at least three independent experiments. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (two tailed t-test).

72 h. The time-dependent experiments showed that **5** reversed *IR* pre-mRNA splicing defects after 72 h incubation, whereas the DB213 control did not (Figure 2.20a and b). More specifically, approximately 60% of splicing defect of *IR* pre-mRNA was corrected in comparison to untreated samples.

The 72-h-incubation time point was selected for dose-dependent experiments in which DM1 model cells were treated with **5** at 20, 50, 75, and 100 μM (Figure 2.20c and d). Treatment DM1 cells with different concentrations of **5** showed a rescue of the *IR* splicing defect at relatively low concentration (20 μM), but no clear dose-dependence result was observed, possibly because splicing is a downstream event which is not directly related to the action of ligands or there might be other factors contributing to the *IR* splicing process. A similar result was obtained for *cTNT* splicing experiments (performed by Jessie Peh).

2.7. Conclusion

Our group has rationally designed small molecules targeting CUG^{exp} based on the crystal structure of the repeat RNA. These simple ligands contain one or two triaminotriazine units designed to target a UU mismatch covalently linked to a known RNA binding moiety (e.g., acridine or bisamidinium unit). Optimized acridine-based and bisamidinium-based ligands inhibited MBNL1-CUG^{exp} complexes *in vitro*, dispersed disease foci, and rescued IR pre-mRNA splicing defect in DM1 model cells. Bisamidinium-based ligands are much less toxic and more water-soluble, whereas the acridine-based ligands generally exhibit stronger binding affinity toward CUG^{exp}. Interestingly, MBNL1-CUG^{exp} inhibition potencies of both sets of ligands were comparable, suggesting that a binding affinity and a physical barrier of ligands should be taken into account to determine how strongly a ligand can inhibit the MBNL1-CUG^{exp} complex. Because of advantages in water solubility, cell uptake, and cytotoxicity, the bisamidinium-based ligand strategy was applied to DM2 and this will be discussed in detail in Chapter 3.

2.8. Materials and methods

MBNL1 Expression and Purification

An expression vector for a truncated MBNL1 comprised of amino acids 1–272 was obtained from Maurice S. Swanson (University of Florida, College of Medicine, Gainesville, FL). This MBNL1 construct is comprised of the four zinc finger motifs of MBNL1 and a hexahistidine tag (C-terminus) and binds RNA with similar affinity as the full length MBNL1. The protein was expressed and purified as described previously (7). The molecular weight was confirmed by MALDI-TOF mass spectrometry, the concentration was determined by Bradford protein assay (Bio-Rad), and the purity determined by silver-stained SDS-PAGE.

Equilibrium Binding Assays

$(CUG)_{12}$ RNA was labeled with [γ-^{32}P]-ATP using T4 poly-nucleotide kinase (New England Biolabs) and labeled RNA was purified by phenol extraction and ethanol precipitation. Labeled RNA and unlabeled RNA was heated at 95 °C for 5 min, then placed on ice for 10 min and diluted to protein binding buffer (175 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 8), 1.25 mM 2-mercapto-ethanol (BME), 12.5% glycerol, 2mg/mL bovine serum albumin (BSA), 0.1 mg/mL heparin, 0.05% or 0.1% Triton X). For MBNL1-r$(CUG)_{12}$ binding assay, MBNL1 was serially diluted in protein binding buffer and labeled and unlabeled RNAs were added to a final concentration of 0.22 nM (unlabeled RNA:labeled RNA = 10:1). The reaction mixture was incubated at room temperature for 20 min and loaded onto a 6% polyacrylamide gel (80:1) at 4 °C. The gel was run for 1.5 h at 150 V in Tris-borate buffer (pH 8). Gels were dried, developed

overnight in phosphor cassette and visualized on a Molecular Dynamics Storm PhosphorImager. The K_d (1:1 stoichiometry assumption) was obtained by fitting a plot of fraction RNA bound versus protein concentration (Prism) using the equation: Fraction RNA bound = $B_{max} \times [MBNL1]_{total}^{h}/(K_d^{h} + [MBNL1]_{total}^{h})$, where B_{max} is maximum fraction RNA bound, h is Hill slope. The protein concentration in the reaction mixture was in a 10-fold excess over the RNA concentration. The average K_d and standard deviation were obtained from three independent experiments.

Inhibition Assays

Ligands **4**–**8** were dissolved in the protein binding buffer with 10% (v/v) DMSO at the highest concentration (3.4 mM) and the resulting stocks were serially diluted. Pentamidine was dissolved in water. *Para*- and *meta*-H1 were dissolved in DMSO. The inhibition assay was performed with the same procedure as in the K_d determination of the MBNL1-r(CUG)$_{12}$ binding assay except the addition of small molecule to MBNL1-RNA complex. A mixture of RNAs, unlabeled and labeled RNAs with the ratio of 10, and MBNL1 was incubated at room temperature. After 15 min of incubation, the MBNL1-RNA complex solution was added to the ligand solution to give final RNA and MBNL1 concentrations of 0.22 nM and 0.1 μM, respectively. The mixture was incubated 15 min and loaded onto a 6% polyacrylamide gel (80:1) at 4 °C. The gel was run for 1.5 h at 150 V in Tris-borate buffer (pH 8). Gels were dried, developed overnight in phosphor cassette and visualized on a Molecular Dynamics Storm PhosphorImager. The fraction RNA bound versus log[compound] was fit (Prism) using the equation: $y = ax^{b}/(IC_{50}^{b} + x^{b})$ where y =bound RNA fraction, x = ligand concentration, $a = y_{max} - y_{min}$, b = hill slope.

The average IC_{50} and standard deviation were obtained from three independent experiments.

Plasmids

Wild type cTNT, DT960 and DT0 minigenes were obtained from the lab of Thomas Cooper (Baylor College of Medicine). The insulin receptor (IR) minigene was obtained from the lab of Nicholas Webster (University of California, San Diego).

FISH (Fluorescence In Situ Hybridization)

A total of ca. 120,000 HeLa cells were seeded in each well of a 6-well plate on coverslips. After a day, the cells were transfected with 500 ng DMPK−CUG0 or DMPK−CUG960 plasmid and 500 ng GFP-MBNL1 plasmid using Lipofectamine following the manufacturer's protocol at cell confluence of 70−80%. After 4 h, the media was changed and ligand was added to each well at desired concentrations. After 36 h or 48 h, the cells were fixed with 4% PFA then washed five times with 1X PBS. Fixed cells were permeabilized with 0.5% triton X-100 in 1X PBS at room temperature for 5 min. Cells were prewashed with 30% formamide in 2X SSC for 10 min at room temperature. Cells were probed with FISH probe (1 ng μL-1 of Cy3 CAG10 in 30% formamide, 2X SSC, 2 μg mL-1 BSA, 66 μg mL-1 yeast tRNA) for 2 h at 37 °C.

Cells were then washed with 30% formamide in 2X SSC for 30 min at 37 °C, followed by washing with 1X SSC for 30 min at room temperature. The cells were washed twice with 1X PBS and then nuclei were stained with 1 μM To-Pro-3 and washed twice. Cells were mounted onto glass sides with ProLong® Gold. Slides were imaged at

RT by LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. The following table indicates the excitation filters used in these experiments.

Confocal Microscopy

Approximate 120,000 HeLa cells were plated in each well of a 6-well plate onto coverslips maintained in complete growth media [DMEM (4.5g/L glucose) supplemented with L-glutamine and 10% FBS (Gemini)]. Cells were transfected at 70–80% confluence with 1 μg of DT960 plasmid with Lipofectamine 2000 (Invitrogen, Life Technologies) according to the manufacturer's protocol. After 4 h, the transfection media was removed and the compound was added to the complete growth media if being assayed. For time course experiments, cells were treated with 100 μM **5**, **7**, and **8** for 48 h.

Cells were fixed for 10 min at room temperature with 4% PFA and washed 5 times for 10 min in 1X PBS at room temperature. Cells were stored at 4 °C if not probed immediately. For the fluorescence in situ hybridization (FISH) procedure, cells were permeabilized with 0.5% triton X-100 in 1X PBS at room temperature for 5 min. Cells were then prewashed with 30% formamide, 2X SSC for 10 min at room temperature. Cells were then probed for 2 h at 37 °C, with 1 ng/μL of Cy3-CAG_{10} probe (Integrated DNA Technologies) in 30% formamide (Fisher Scientific), 2X SSC, 20 μg/mL BSA, 66 μg/mL yeast tRNA (Sigma-Aldrich). Cells were then washed for 30 min in 30% formamide, 2X SSC at 37 °C, and then with 1X SSC for 30 min at room temperature.

For immunofluorescent (IF) staining, cells were washed twice in 1X PBS, 10 min at room temperature, and then probed overnight at 4 °C with anti-MBNL1 antibody (1:5000 dilution, clone 3A4 antibody, Millipore) in 1X PBS. Cells were washed 2 times for 10 min at room temperature with 1X PBS. Next, cells were incubated with goat anti-mouse AlexaFluor 488 antibody (1:500 dilution, Invitrogen) for 2 h at room temperature. Cells were then washed 2 times for 10 min at room temperature with 1X PBS and then treated with 10 μg/mL Hoechst 33342 (Sigma-Aldrich) for 5 min. Finally, cells were washed with 1X PBS for 5 min and mounted onto glass slides with fluorescent mounting medium (Dako) before sealing with nail polish. Cells were imaged on Zeiss confocal laser scanning microscope (LSM) 700 at 20× or 40× magnification.

Live cell imaging

A total of ca. 120,000 HeLa cells were grown in an Ibidi 35 mm Petri dish with a standard bottom, high walls and an imprinted 500 Pm relocation grid. After a day, cell confluence reached to about 70−80%, cells were tranfected with 500 ng DMPK−CUG960 plasmid and 500 ng GFP-MBNL1 plasmid using Lipofectamine following standard protocol. After 4 h, media were changed and cells were incubated at 37 °C, 5% CO2. 24 h post-transfection, ligand was added to a desired concentration. Live-cell, time-lapse images were taken before addition of 1 as well as at 2, 4 and 7 h time points at RT by a LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. For tracking the cells, DIC images were acquired simultaneously with the reflected light images using a TPMT

module after setting the Köhler illumination with a fully opened condenser aperture (0.55 NA).

Splicing assays

Approximately 120,000 HeLa cells were seeded in each well of a 6-well plate in complete growth media the day before transfection. For testing the IR pre-mRNA, cells were similarly transfected the following day at 70–80% confluence with 500 ng each of IR and DT960 or DT0 minigenes with Lipofectamine 2000 (Invitrogen, Life Technologies) according to the manufacturer's protocol. After 4 h, the transfection media was removed and ligand was added to the complete growth media if being assayed. Cells treated with 75 μM of pentamidine were harvested 16–20 h post-treatment because of observed toxicity. For time course experiments, cells were treated with 100 μM of ligand, harvested after 24, 48, or 72 h for **5** and after 72 h for DB213. Cells were treated with **3** at concentrations of 20, 50, 75 and 100 μM, harvested after 72 h. To harvest the cells, cells were washed once with 1X PBS and detached using trypsin with 0.05% EDTA (Fisher Mediatech). Following harvesting, RNA was immediately isolated using Total RNA kit I (Omega Bio-Tek). 900 ng of isolated RNA was reverse transcribed with iScript cDNA synthesis kit (Bio-Rad) according the manufacturer's protocol. The reverse transcription reaction was cleaned up using QIAquick PCR purification kit (Qiagen). Approximately 70 ng of cDNA was subjected to 31–35 cycles of PCR amplification using gene specific primers. The forward primer was 5'-GTA CCA GCT TGA ATG CTG CTC CT, and the reverse primer was 5'-CTC GAG CGT GGG CAC GCT. The linear range for PCR for the IR constructs was found to be between 25–35 cycles. The PCR products were ran on

an 8% polyacrylamide gel with 1X TBE (National Diagnostics) at 120 V for 55 min. The gel was post stained with ethidium bromide and subsequently imaged using Gel Doc XR+ system (Bio-Rad). The bands were quantified using ImageJ (NIH).

SRB Toxicity Assay

The SRB assay was performed according to the method of Vichai and Kirtikara.[7] DMEM supplemented with 10% FBS (98 μL) was placed in a 96-well plate. Ligand solution (2 μL) was added to give final concentrations from 5 nM to 100 μM, five repeats for each concentration. HeLa cells (10,000 cells/well), DM1 fibroblast cells (GM03132, ATCC) (10,000 cells/well), 3T3 cells (ATCC) (10,000 cells/well) were then plated in the 96-well plate. The cells were incubated at 37 °C. After 72 h, 100 μL of cold 10% (w/v) trichloroacetic acid were added in each well, followed by incubation at 4 °C for 1 h. The cells were washed twice with tap water and then air-dried. The cells were stained with 100 μL of 0.0057% (w/v) sulforhodamine B in 1% acetic acid at room temperature, 30 min. The plate was rinsed twice with 1% (v/v) acetic acid to remove the unbound stain. The bound protein stain was solubilized in 200 μL of 10 mM Tris base, pH 10.5 and left for 30 min. The optical intensity was measured at 510 nm using microplate reader (SPECTRAmax PLUS). Percentage of dead cells was calculated using the following formula:

$$\%\ \text{cell dead} = 100\% - \frac{\text{Mean}_{\text{OD sample}} - \text{Mean}_{\text{OD dead control}}}{\text{Mean}_{\text{OD live control}} - \text{Mean}_{\text{OD dead control}}} \times 100\%$$

The data was plotted using Excel and curves were fitted using Table Curve (Systat) to estimate the IC_{50} values. The cytotoxicity experiment was triplicated for HeLa cells.

2.9. Reference

(1) Childs-Disney, J. L.; Stepniak-Konieczna, E.; Tran, T.; Yildirim, I.; Park, H.; Chen, C. Z.; Hoskins, J.; Southall, N.; Marugan, J. J.; Patnaik, S.; Zheng, W.; Austin, C. P.; Schatz, G. C.; Sobczak, K.; Thornton, C. A.; Disney, M. D. *Nat. Commun.* **2013**, *4*, 1–11.

(2) Pinheiro, P. *J. Biol. Chem.* **2002**, *277*, 35183–35190.

(3) Mooers, B. H. M.; Logue, J. S.; Berglund, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16626–16631.

(4) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. *Nucl. Acids Res.* **2009**, *37*, 4149–4156.

(5) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16068–16073.

(6) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. *ACS Chem. Biol.* **2013**, *8*, 1037–1043.

(7) Vichai, V.; Kirtikara, K. *Nat. Protoc.* **2006**, *1*, 1112–1116.

(8) Schoser, B.; Timchenko, L. *Current genomics* **2010**, *11*, 77.

(9) Savkur, R. S.; Philips, A. V.; Cooper, T. A. *Nat. Genet.* **2001**, *29*, 40–47.

(10) Hung, M.; Patel, P.; Davis, S.; Green, S. R. *J. Virol.* **1998**, *72*, 4819–4824.

(11) Marcheschi, R. J.; Tonelli, M.; Kumar, A.; Butcher, S. E. *ACS Chem. Biol.* **2011**, *6*, 857–864.

(12) Ofori, L. O.; Hoskins, J.; Nakamori, M.; Thornton, C. A.; Miller, B. L. *Nucl. Acids Res.* **2012**, *40*, 6380–6390.

(13) Warf, M. B.; Nakamori, M.; Matthys, C. M.; Thornton, C. A.; Berglund, J. A.

Proc. Natl. Acad. Sci. U.S.A. **2009**, *106*, 18551–18556.

(14) Parkesh, R.; Childs-Disney, J. L.; Nakamori, M.; Kumar, A.; Wang, E.; Wang, T.; Hoskins, J.; Tran, T.; Housman, D.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 4731–4742.

(15) Wong, C.-H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 6355–6361.

Chapter 3

Development of novel CCUG^{exp}-MBNL1 inhibitors[1]

3.1. Introduction

Myotonic dystrophy type 2 (DM2) is associated with an expansion of CCTG repeats (CCTG^{exp}) in the intron 1 of the *ZNF9* gene on chromosome 3.[1] In comparison with DM1, DM2 shows less severe symptoms with the estimated prevalence of 1 in 10000 people.[2] The minimum length of the repeats indicating the presence DM2 diseases is 75 CCTG repeats. The sequence is unstable, progressively expanding and reaching thousands of CCTG repeats in the most severe cases of the disease. Transcribed CCTG repeats (CCUG^{exp}) sequester muscleblind-like proteins including an important alternative-splicing regulator, MBNL1. A decrease in cellular MBNL1 levels results in splicing defects of >100 pre-mRNAs, leading to the disease symptoms, which include myotonia, muscle weakness, and cataracts.[2] In addition, the CCUG transcript appears to interact with other proteins that are involved in translation.[2-4] Thus, it was reported that the CUG binding protein (CELF1) and a translation initiation factor (eIF2) form a complex and both are sequestered by CCUG^{exp} as are the proteasome, both events leading to a reduction of the global rate of protein synthesis.[3-5] The CCUG^{exp} transcript may affect the expression levels of the ZNF9 protein, and this may contribute to the DM2 phenotype, however the reports on this point are somewhat contradictory.[4-8] An important potential approach to treating DM2 involves molecular intervention, the discovery and development of either oligonucleotides or small molecules that selectively disrupt the interaction of CCUG^{exp} with MBNL1 and other proteins. This same general approach has been applied to DM1 and is considerably more advanced.[9-19] To our knowledge there are only two reports of small molecules that selectively target the toxic

[1] The material in this chapter was adapted from the following publication:
Nguyen, L.; Lee, J.; Wong, C.-H.; Zimmerman, S. C. *ChemMedChem* **2014**, *9*, 2455-2462.

CCUG^{exp} transcript causing DM2 (Figure 3.1).[20,21] In 2009, Disney and coworkers described a series of multivalent ligands displaying kanamycin A units with varying separations (see 1, n =

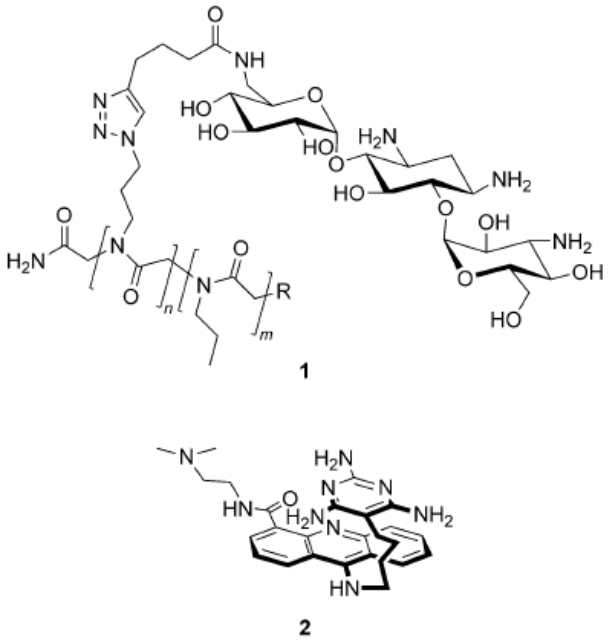


Figure 3.1. Two reported small molecules targeting DM2

1-3, m = 0-19), the best was the trimer which inhibited the $(CCUG)_{47}$-MBNL1 interaction with nanomolar IC_{50} values in vitro.[20] In 2011, and as described in more detail below, our group reported ligand **2** as a strong and selective inhibitor of the $(CCUG)_6$-MBNL1 interaction.

3.2. Design of groove binding DM2 ligands

The design of novel small molecules targeting CCUG^{exp} was based on the knowledge gained in recent work on ligands **2**[21] and **3**[16] that target DM1 and DM2 (Figure 3.2). Our group developed a simple ligand **2** for DM2 based on DM1-targeting ligand **3**. Ligand **2** was composed of a triaminopyrimidine moiety recognizing CU mismatches in CCUG^{exp} hairpin and an acridine unit. The protonation of the pyrimidine due to its higher pK_a leads to a Janus wedge moiety that was proposed to complement with the CU mismatch.[21] The selectivity of ligands **2** and **3** was shown toward CUG^{exp} and CCUG^{exp}, respectively, despite the structural similarities of the

triaminotriazine and triaminopyrimidine moieties. Thus, the triaminopyrimidine unit is key to our design of novel DM2 ligands.

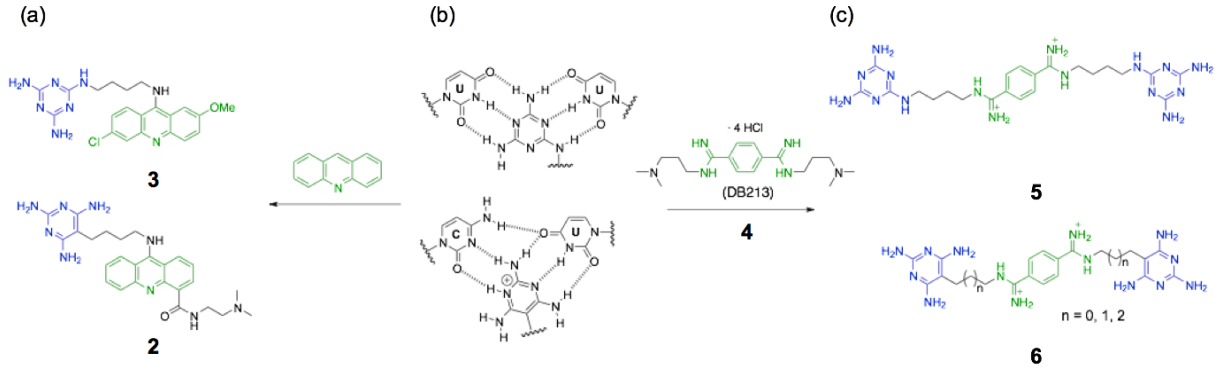


Figure 3.2. Design of bisamidinium-based ligands targeting DM2. (a) Acridine-based ligands for DM1. (b) Proposed binding modes of a triaminotriazine and a triaminopyrimidine to UU and CC mismatches, respectively. (c) Bisamidinium-based ligands targeting DM1 and DM2.

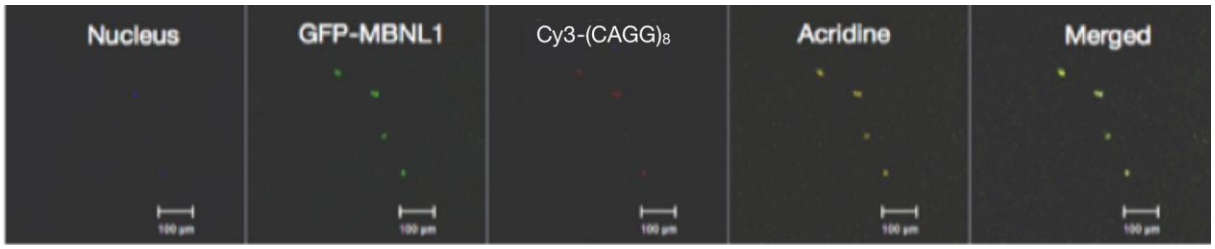


Figure 3.3. Confocal image of cells treated with ligand 2. DM2 model cells as HeLa cells co-transfected with plasmids expressing $(CCTG)_{1200}$ and GFP-MBNL1 were treated with 2 (50 μM, 48 h). $CCUG^{exp}$ was stained by a FISH probe, Cy3-$(CAGG)_8$.

Ligands **2** and **3** contain an acridine intercalator that provides a hydrophobic driving force for binding. In addition, the stacking of the intercalating unit on the Janus wedge unit was expected, which showed a success in the suppression of the off-target binding of duplex DNA of ligand **3**. However, ligands **2** and **3** exhibited poor water solubility, low cell uptake, and high toxicity. For example, treatment of HeLa cells with ligand **2** resulted in the death of more than 90% of DM2 model cells after 48 h incubation (Figure 3.3). Bisamidinium compounds are known to possess excellent water solubility, high cellular uptake,[22] and low cytotoxicity.[23] In

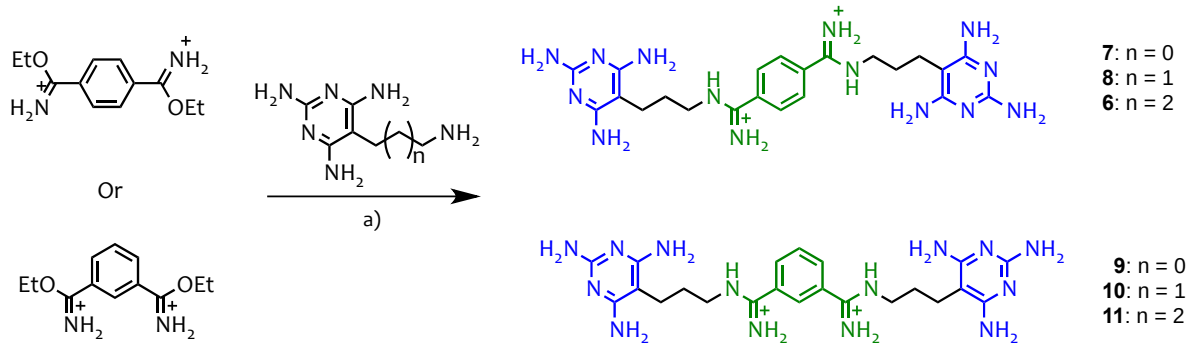
addition, bisamidiniums with the appropriate substituents are taken up by the nucleus.[22] Thus, ligand **4**, a groove-binding bisamidinium ion, was selected to replace the acridine unit in the novel DM2 small molecules. In 2011, Butcher and co-workers reported the NMR structure of ligand **4** (DB213, R = NMe_2) bound within the major groove of the HIV-1 frameshift RNA.[24] Importantly, Wong and coworkers[19] noted that the HIV-1 stem-loop RNA adopts a remarkably similar A-form conformation to $(CUG)_{12}$, and the distance between the ammonium ions in bound **4** is the same as between every third UU mismatch in $(CUG)_{12}$.

As a conjugate of the bisamidinium groove binder **4** and two triaminotriazine recognizing UU mismatches, ligand **5** exhibited relatively low toxicity in HeLa, mouse myoblast, and DM1 fibroblast, and was well tolerated by mice. Importantly, **5** was found to bind CUG^{exp}, disrupt the disease foci, rescue cTNT and IR pre-mRNA splicing defects in DM1 model cells, and improve the eye phenotype in a DM1 *Drosophila* model.[19] Whether $CCUG^{exp}$ has a structural similarity to CUG^{exp} would determine if the similar approach could be applicable for DM2. It was predicted that $CCUG^{exp}$ might adopt two possible structures: one with two consecutive CU mismatches or one referred as the slipped form with alternating CC and UU mismatches.[21,25] After this work was completed, in 2014, Disney and co-workers reported the first high-resolution X-ray structure of an $(CCUG)_3$ duplex within a larger RNA sequence. showing that the CCUG sequence indeed adopts an A-form structure with two CU mismatches separated by two GC base pairs.[26]

3.3. Synthesis of bisamidinium-based DM2 ligands

With the considerations outlined above, analogous bisamidinium ligands for DM2 were synthesized by conjugating the bisamidinium groove binding unit and two triaminiopyrimidine Janus wedge units that target the CU mismatches in $CCUG^{exp}$ (Figure 3.2).[21] The substitution pattern of the bisamidinium unit may affect its cell permeability and localization, and its specific

fit to CCUG[exp]. Thus, ligands with para- and meta-substituted bisamidinium moieties were prepared, in which alkyl linkers vary from two to four methylene groups. This set of ligands was used to determine the optimum length and general fit in the groove.



Scheme 3.1. Synthesis of **6-11**. Reagents and conditions: a) Et_3N, EtOH, 25 °C, 24 h, 39-52%

Ligands **6-11** were prepared from the appropriate meta- or para-bisimidate and the corresponding aminoalkylpyrimidine (Scheme 3.1 and Materials and Methods). The aminoalkylpyrimidine compounds were prepared using a modification of the procedure reported by Mascal and co-workers.[27] Briefly, the dicyanoalkyl phthalimide reacted with guanidine carbonate, followed by deprotection using hydrazine (Scheme 3.2 and Materials and Methods).

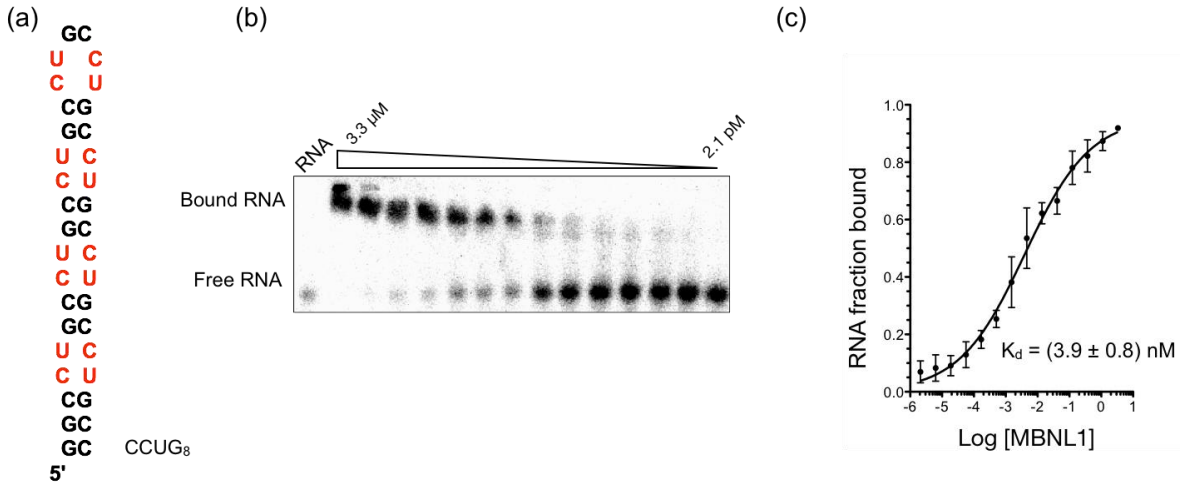


Figure 3.4. Binding study of $(CCUG)_8$ and MBNL1. (a) Structure of $(CCUG)_8$. (b) Binding gel and (c) binding curve of $(CCUG)_8$ and MBNL1. Error bars represent standard deviations from three independent experiments. For experimental details, see Materials and methods section at the end of this chapter.

3.4. Inhibition of MBNL1-$(CCUG)_8$ complex by DM2 small molecules *in vitro*

Electrophoretic mobility shift assay (EMSA) was performed to determine the inhibition potency of DM2 ligands toward MBNL1-$(CCUG)_8$. The complementary GC-rich regions at the 3'- and 5-ends were used to avoid formation of the slipped $CCUG^{exp}$ and restrict the structure to that with two consecutive CU mismatches (Figure 3.4a). The binding affinity of MBNL1N for $(CCUG)_8$ was determined using EMSA. Thus, radiolabeled $(CCUG)_8$ was incubated with serially diluted MBNL1N, and the relative amount of free and complexed RNA were measured after a gel separation. The percent RNA bound by MBNL1N was plotted against the protein concentration, and curve fitting provided a K_D value of 3.9 ± 0.8 nM (Figure 3.4b and c).

The inhibition potency of DM2 ligands **2**, **6-11** and DM1 ligands **5**, **12-14** and the control compound **4** were qualitatively screened by incubating each at 100 µM with MBNL1N-$(CCUG)_8$ complexes. The relative ratios of free and MBNL1N-complexed $(CCUG)_8$ were measured (Figure 3.5).

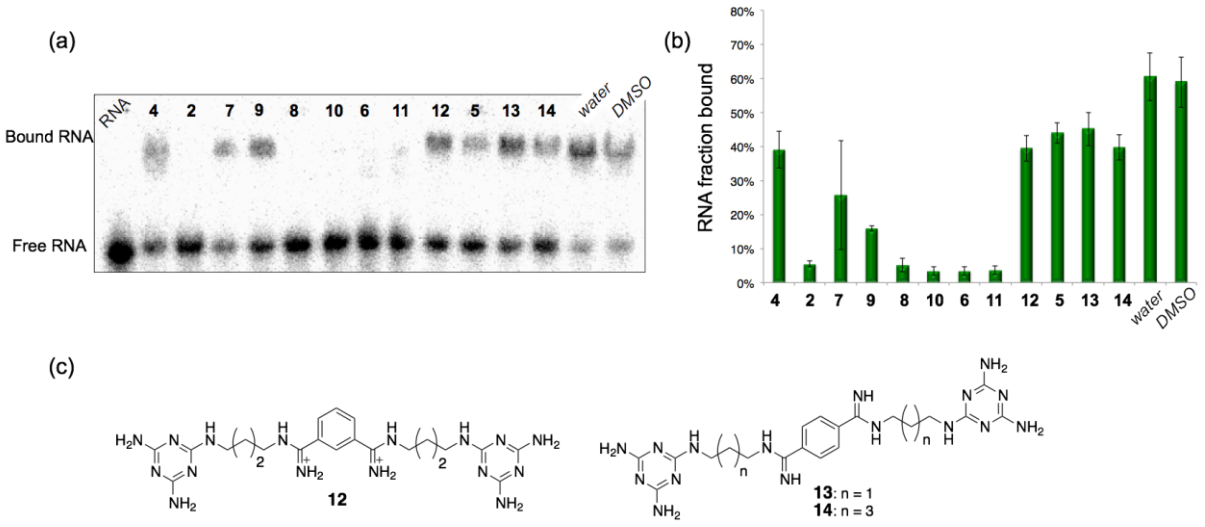


Figure 3.5. Screening experiment. (a) Screening gel. In this gel, [$(CCUG)_8$]= 0.22 nM, [MBNL1]= 33 nM, [ligand]=100 µM. (b) Graph of the percentage of RNA fraction bound versus ligands, showing **6**, **8**, **10**, and **11** were the potent ligands. Error bars represent standard deviations from three independent experiments (c) Structure of ligand included in the screening assay. For experimental details, see Materials and methods section at the end of this chapter.

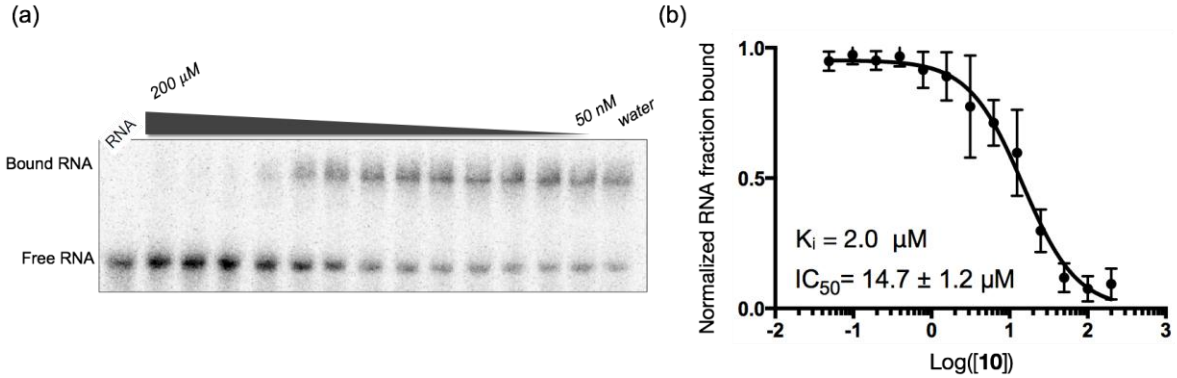


Figure 3.6. EMSA inhibition experiment. (a) Inhibition gel of ligand **10**. Briefly, MBNL1– $(CCUG)_8$ complex was incubated with **10** at different concentrations (200 μM to 50 nM, twofold dilution). (b) The inhibition curve of ligand **10**; the RNA fraction bound was normalized to the control. Error bars represent the standard deviation of at least three independent experiments. For more experimental details, see Materials and methods section at the end of this chapter.

Table 3.1. Inhibition potency of DM2 ligands to MBNL1- $(CCUG)_8$.

Compound	K_D (nM) (MBNL1- $(CCUG)_8$)	Half-maximal inhibitory concentration IC_{50} (μM)	Inhibition constant K_i (μM)
6	3.9 ± 0.8	73.4 ± 6.0	8.7 ± 1.9
8		12.9 ± 1.0	1.5 ± 0.3
10		14.7 ± 1.2	1.7 ± 0.4
11		24.9 ± 2.0	2.9 ± 0.6

Data represent the mean ± SD of three independent experiments performed in duplicate

Compounds **6**, **8**, **10**, and **11** that contain linkers with three and four methylene groups fully disrupted the MBNL1N-$(CCUG)_8$. In contrast, there was a significant amount of complex remained after incubation with the shorter ligands, **7** and **9**, as well as for the negative control **4** (DB213, R = NMe_2) lacking triaminopyrimidine recognition units. Likewise, the DM1 targeting ligands containing triaminotriazine units failed to inhibit the MBNL1N-$(CCUG)_8$ complex (Figure 3.5). Given the results of the screening assay, in collaboration with JuYeon Lee, we performed EMSA inhibition assays to determine IC_{50} values for ligands **6**, **8**, **10**, and **11** (Figure

3.6). The corresponding inhibition constants K_i were in the low micromolar range (Table 3.1). Interestingly, there is no significant difference in inhibition potency between meta- and para-bisamidinium ligands. This result was unexpected because the substitution pattern changes both the distance that the recognition units can span as well as the fit of the bisamidinium unit within the RNA groove. Additional structural data is needed to explain the insensitivity to the benzamidinium substitution pattern.

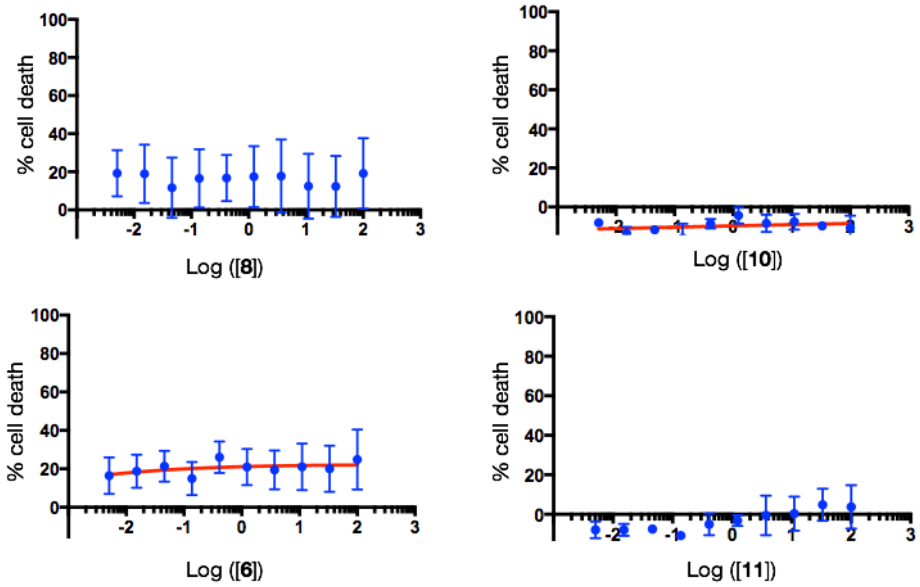


Figure 3.7. Toxicity curves for ligands **6**, **8**, **10**, and **11**. The IC_{50} values for the toxicity, the concentration of ligand causing 50% cell death, was estimated from the plots. The error bars represent standard deviations from three independent experiments. For more experimental details, see Materials and methods section at the end of this chapter.

3.5. Low toxicity of DM2 ligands in HeLa cells

Development of the acridine-based ligand **2** was limited by its high cytotoxicity. In particular, it was found that treatment of DM2 model cells as HeLa cells transfected with plasmids containing $(CCTG)_{1200}$ with **2** at 50 µM led to more than 90% cell death, a change in cell morphology, and no significant reduction in the number disease foci after 48 h (Figure 3.3). In contrast, ligands **6**, **8**, **10**, and **11** showed a low cytotoxicity at 100 µM as incubated with DM2 model cells.

In collaboration with JuYeon Lee, cytotoxicity of bisamidinium-based DM2 ligands were quantitatively determined using a sulforhodamine B (SRB) colorimetric assay[28] with HeLa cells. The percentage of cell death was plotted against ligand concentration (Figure 3.7), giving IC_{50} values from curve fitting. Bisamidinium conjugates showed a low cytotoxicity with an IC_{50} of more than 100 μM for 3 d incubation, which is a significant improvement in comparison with ligand **2**. In particular, para-bisamidinium based ligands **6** and **8** were observed to have low toxicity with between 15 and 30% cell death, whereas the meta-bisamidinium compounds **10** and **11** exhibited essential no cytotoxicity under the testing conditions.

3.6. Foci disruption in DM2 model cells by small molecules

The hallmark of DM2 disease is the formation of foci resulting from the sequestration of MBNL proteins by $CCUG^{exp}$. A co-localization of MBNL1 protein with $CCUG^{exp}$ in DM2 patient cells was clearly observed as punctate foci in immunofluorescence experiments.[29] With a similar pathogenesis to DM1, it is believed that the DM2 phenotypes might be reversed by small molecules able to disrupt the MBNL1-$CCUG^{exp}$ complex in cells. As mentioned in the introduction there was a ligand reported by Disney to rescue the splicing defect, a downstream event, in DM2 model cells. However, no direct observation of small molecules dissolving disease foci in DM2 cell culture has been reported. Importantly, treatment of DM1 model cells with ligands targeting toxic CUG^{exp} repeats and interrupting the MBNL-CUG^{exp} interaction resulted in the reversal of the DM1 phenotypes.[9-19]

To examine whether the *in vitro* activity of **6**, **8**, **10**, and **11** translated into cellular activity, fluorescence *in situ* hybridization (FISH) experiments were performed to study the ability to dissolve foci in DM2 of lead ligands **6**, **8**, **10**, and **11**. Thus, HeLa cells were co-transfected with GFP-MBNL1 and $(CCTG)_{1200}$ plasmids to generate DM2 model cells.[3,4,26] The locations of

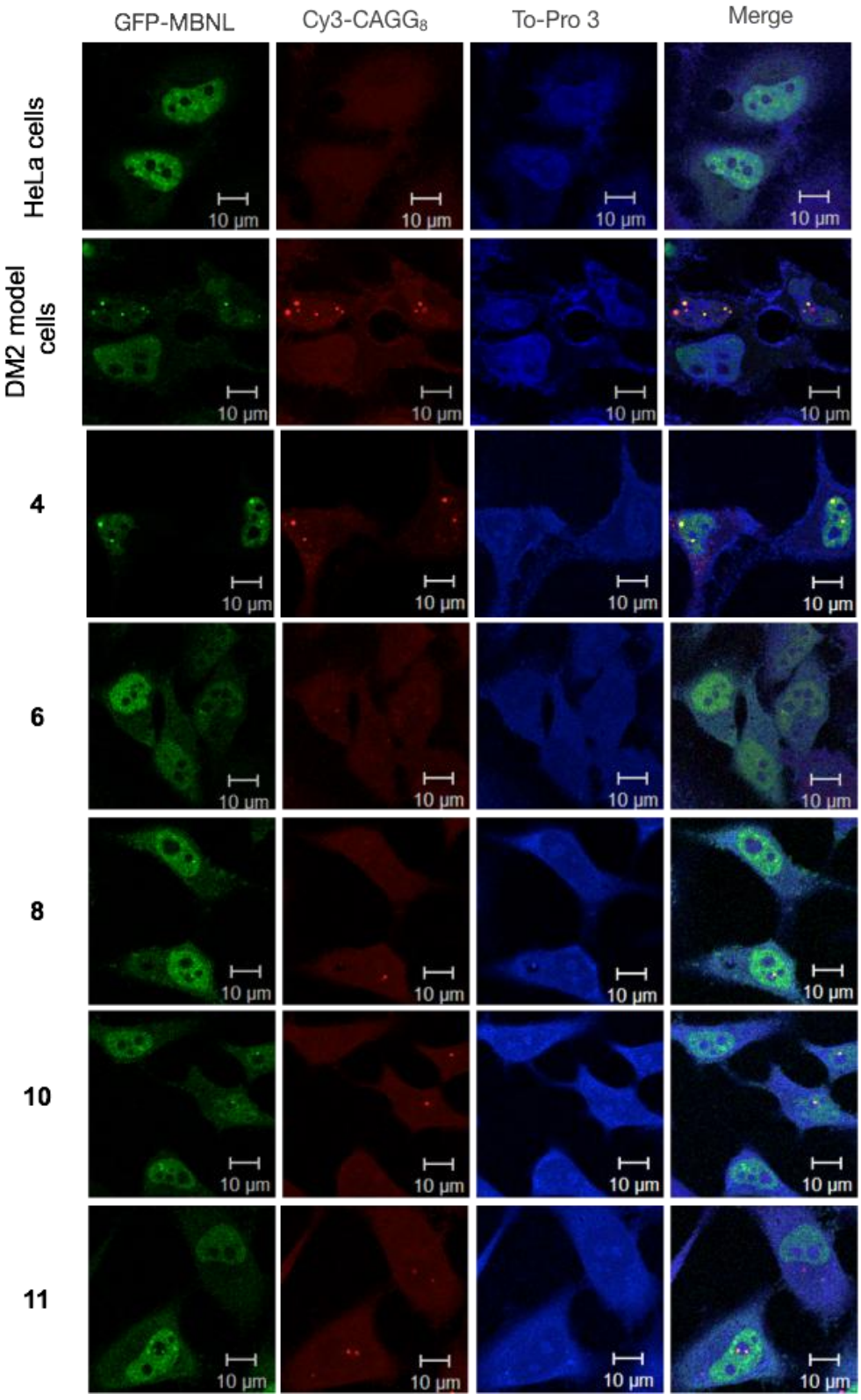


Figure 3.8. Confocal microscopy images of foci dispersion. DM2 model cells were treated with compounds **4**, **6**, **8**, **10**, **11** at 100 µm, 48 h; r(CCUG)n was probed by Cy3-$(CAGG)_8$; the nucleus was stained by To-Pro 3. For experimental details, see Materials and methods section at the end of this chapter.

MBNL1 and the $(CCUG)_{1200}$ transcripts were readily tracked by confocal microscopy, monitoring the fluorescent signal of GFP and Cy3-$(CCAG)_8$ FISH probe, respectively (Figure 3.8).

The CCUG repeats were found to localize in both nuclei and cytoplasm, which is consistent with a report by Timchenko and coworkers.[4] A co-localization of GFP-MBNL1 and $(CCUG)_{1200}$ within the disease foci was clearly observed, providing evidence of MBNL1 sequestration by CCUG repeats (row 2 in Figure 3.8). In addition, there were some $(CCUG)_{1200}$ foci not co-

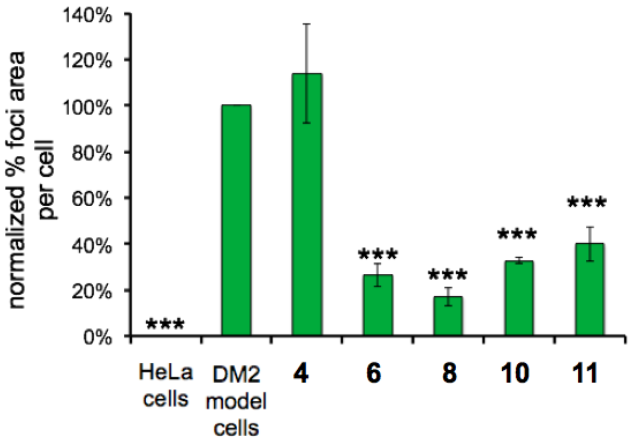


Figure 3.9. Plotting percent normalized foci area per against different treatment conditions. DM2 model cells were treated with compounds **4**, **6**, **8**, **10**, **11** at 100 µm. Error bars represent the standard error of the mean of at least three independent experiments. *** indicates a statistically significant difference in comparison to untreated DM2 model cells, $P < 0.001$ (two tailed *t*-test). For more experimental details, see Materials and methods section at the end of this chapter.

localized with GFP-MBNL1, which might suggest a limited co-transfection of the protein or possibly the $(CCUG)_{1200}$ transcript bound by endogenous MBNL or other protein complexes including the 20S catalytic core and the CELF1-eIF2 complex.[3,4]

All of the bisamidinium ligands that inhibited the MBNL1-$(CCUG)_8$ complex *in vitro* dissolved the disease foci with no co-localization of Cy3 and GFP signals (rows 4-7 in Figure 3.8) whereas the control ligand **4** (DB213, R=NMe_2) showed almost no foci dispersion (row 3 in

Figure 3.8). Treatment of DM2 model cells with ligands **6**, **8**, **10**, and **11** resulted in a significant decrease in the number and size of the foci. Thus, 50-70% foci disruption was observed after 48 h treatment at 100 μM (Figure 3.9). There was no correlation between the K_i and the extent of foci disruption, which is understandable given that many other factors govern the cellular activity.

3.7. Conclusion

One therapeutic approach for myotonic dystrophy uses small molecules to target CUG or CCUG repeats and disrupt their interaction with MBNL1 and other proteins. In this chapter, a new set of ligands that inhibited the MBNL1-$CCUG^{exp}$ interaction was reported. The ligands with low molecular weight of approximately 520 Da contain two important units: a triaminopyrimidine moiety for targeting the CU mismatch and a bisamidinium moiety. The bisamidinium unit provides a general A-form RNA targeting ability, confers excellent water solubility, and cellular and nuclear uptake. The linker length between the triaminopyrimidine recognition and the bisamidinium units and substitution pattern on the benzene ring were optimized. Four in six ligands show promising biological activity, inhibiting MBNL1-$(CCUG)_8$ with low micromolar K_i values, very low cytotoxicity in HeLa cells, and dissolving the disease foci in DM2 model cells. Indeed, the low cytotoxicity of ligands **6**, **8**, **10**, **11** and the binding selectivity of ligand **2** toward a sequence containing CU mismatches[21] suggest a specificity of bisamidinium ligands toward $CCUG^{exp}$. With the promising results from *in vitro* inhibition experiments and cellular study, the bisamidinium-based DM2 ligands are appropriate for further testing to their ability of rescuing the disease phenotypes in DM2 *Drosophila* and mouse models.

3.8. Materials and methods

Biological experiments

Materials: r(CCUG)$_8$ (5'-GCCCUGCCUGCCUGCCUGCCUGCCUGCCUGCCUGGC-3') was purchased from Dharmacon RNAi Technologies (GE Healthcare) with 2'-deprotection, desalting, and HPLC purification. The r(CCUG)$_8$ was dissolved in Ambion (Life Technologies) RNA storage solution (pH 6.4) to give a 1 mM stock solution.

Protein expression and purification: An expression vector for a truncated MBNL1 contains amino acids 1–272. This MBNL is comprised of the four zinc-finger motifs of MBNL1and a His6, which has a similar affinity as the full-length MBNL1. The protein was expressed and purified following the procedure described previously.[29]

RNA radioactive labeling: (CCUG)8 was labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase (New England Biolabs) and purified by phenol extraction and ethanol precipitation. Labeled RNA was heated at 95 ºC for 5 min, placed on ice for 10 min, and diluted in a binding buffer (175 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris (pH 7.5), 1.25 mM 2-mercaptoethanol, 12.5% glycerol (v/v), 2 mg mL^{-1} bovine serum albumin (BSA), 0.1 mg mL^{-1} heparin, 0.05% Triton-X (v/v)).

Binding experiments: (CCUG)$_8$ at a final concentration of 0.22 nm was incubated with MBNL1 at serially diluted concentrations (3.3 μM to 2.1 pM). The reaction mixture was incubated at RT for 15 min, and loaded on 6% polyacrylamide gel (80:1) at 4 ºC. The gels were run for 2 h at 200 V in 0.5M Tris borate buffer (pH 8.2). Gels were visualized on a Molecular Dynamics Storm

PhosphorImager. The apparent K_D was obtained by fitting fraction RNA bound versus protein concentration using the equation: Fraction RNA bound = $1/(1+K_D/[MBNL1])$. The standard deviation was obtained from three independent experiments.

Screening EMSA experiments: Each ligand at 100 μM was incubated with pre-incubated MBNL1–$(CCUG)_8$ complexes at RT for 15 min. Final concentrations of MBNL1 and $(CCUG)_8$ were 33 nM and 0.22 nM, respectively. The reaction mixture was loaded on the gel as described above. Fraction RNA bound was compared to the control samples (H_2O or DMSO) to determine the hits under the testing conditions.

Inhibition EMSA experiments: Inhibition experiments were done as described above except MBNL1–$(CCUG)_8$ complexes were incubated with ligands to give a serially diluted concentrations (from 200 μM to 50 nM). The IC_{50} values were obtained by fitting fraction RNA bound versus the log function of ligand concentrations using the equation: $Y=B+(A-B)/(1+10^{(X-logIC50))}$, in which Y = fraction RNA bound, B = minimum fraction RNA bound, A=maximum fraction RNA bound, X = log [ligand]. The inhibition constants (K_i) were obtained using the equation: $Ki = IC_{50} \times K_{D((CCUG)8\text{-}MBNL1)}/ [MBNL1]$. Standard deviations were obtained from three independent experiments.

Cytotoxicity study: Toxicity of ligands in HeLa cells was studied following the protocol reported by Vichai and Kirtikara.[28] Briefly, in a 96-well plate, HeLa cells were incubated at 37 ºC for 72 h with the treatment of the ligand. The ligand was added to each well to have serially diluted concentrations from 100 μM to 5.1 nM. The cells were fixed after 3 d and stained with

sulforhodamine B (SRB). Bound SRB was dissolved by Tris base solution (pH 10.5) and the OD was measured at 510 nm. IC_{50} values of toxicity were obtained by plotting the percentage of cell death against concentrations of ligands.

Fluorescence in situ hybridization (FISH): FISH experiments were performed following the protocol described previously.[15] Briefly, HeLa cells at 80% confluence were co-transfected with $(CCTG)_{1200}$ and GFP-MBNL1 plasmids. Transfected HeLa cells were treated with ligands **4**, **6**, **8**, **10**, and **11** at 100 μM, 48 h. rCCUG repeats were probed with Cy3-(CAGG)8, and the nuclei were stained with To- Pro 3. The cells were looked at under confocal microscope LSM 700 (Zeiss). GFP, Cy3 and To-Pro 3 were excited at 488, 564, and 639 nm, respectively. Four random areas were selected to determine the average foci area per cell, using Teton software (version 4.7). The average foci area per cell of untreated cells was set to 100%. Each experiment was repeated at least three times independently.

Synthesis of ligands tested in the study

Triaminopyrimidine–bisamidinium conjugates were synthesized following Scheme S1 in the Supporting Information. ^{1}H NMR spectra were recorded on a Varian Unity Inova 500 spectrometer. All NMR measurements were carried out in [d_6]-DMSO at RT unless otherwise stated. Chemical shifts are reported as parts per million (ppm). Mass spectra were obtained in the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois, Urbana-Champaign, USA. The analytical HPLC traces were obtained using a C18 column with 0.1% TFA in 1:1 (v/v) MeOH/H_2O as eluent. All reactions were carried out under a nitrogen atmosphere. All solvents and reagents were purchased commercially with a reagent quality and

used without further purification. The progress of reaction was monitored by thin-layer chromatography (TLC) using Merck pre-coated silica gel $60F_{254}$.

Ligand **2** was prepared using the procedure reported by Wong and co-workers, the characterization data was identical to that reported.[citation] The synthesis of intermediates can be found in the Supporting Information. HPLC traces of triaminopyrimidine–bisamidine conjugates show a purity of $\geq 95\%$ (see Supporting Information).

***N^1,N^4-Bis(4-(2,4,6-triaminopyrimidin-5-yl)butyl)terephthalimidamide hydrochloride sal*t** (**6**): Et3 N (0.14 mL, 1.0 mmol) was added to diethyl terephthalimidate dihydrochloride (150 mg, 0.5 mmol) in EtOH (10 mL) cooled in an ice bath. A solution of 5-(4-aminobutyl)- pyrimidine-2,4,6-triamine (215 mg, 1.1 mmol) in EtOH (3 mL) and ethylene glycol (3 mL) was added dropwise. The mixture was warmed to RT and stirred for 24 h. The reaction was monitored by TLC using 7:3 H_2O/HOAc as eluent. The solvent was removed *in vacuo*. The residue was dissolved in MeOH and loaded onto a silica gel flash column and purified by chromatography ($MeOH/CH_2Cl_2$ = 1:4 to 4:1). The starting material and a side product were eluted, and the desired product was eluted with 4:1 $MeOH/CH_2Cl_2$ containing 0.04–0.08 mL of a 4M aq HCl solution in 1,4-dioxane per liter of solvent. The desired product fractions were combined, and the solvent was removed *in vacuo* to afford 6 as a light yellow solid, a hydrochloride salt (160 mg, 48%): mp: > 230 ºC (decomp.); ^{1}H NMR (500 MHz): δ = 10.36 (br s, NH_2, 2H), 9.79 (br s, NH_2, 2H), 9.45 (br s, NH_2, 2H), 7.99 (s, ArH, 4H), 7.15 (br s, NH_2, 12H), 3.49 (br s, CH_2, 4H), 2.33 (br s, CH_2, 4H), 1.75 (br s, CH_2, 4H), 1.39 ppm (br s, CH_2, 4H); MS (ESI): m/z (%): 521.1 (100) $[M+H]^+$.

N^1,N^4-Bis(2-(2,4,6-triaminopyrimidin-5-yl)ethyl)terephthalimidamide hydrochloride salt (**7**): The desired compound was prepared using a procedure analogous to that described for 6 using diethyl terephthalimidate dihydrochloride (217 mg, 0.7 mmol) in EtOH (5 mL), Et_3N (0.2 mL, 1.4 mmol), and 5-(2-aminoethyl)pyrimidine-2,4,6-triamine (309 mg, 1.8 mmol) to afford of **7** as a light yellow solid (240 mg, 52 %): ^{1}H NMR (500 MHz): δ = 10.37 (br s, N*H*$_2$, 2 H), 9.75 (br s, N*H*$_2$, 2 H), 9.64 (br s, N*H*$_2$, 2 H), 8.04 (s, Ar*H*, 4H), 6.85 (br s, N*H*$_2$, 8H), 6.54 ((br s, N*H*$_2$, 4H), 3.49 (br t, J=2.5, C*H*$_2$, 4H), 2.71 (br t, J=7.5, C*H*$_2$, 4H); MS (ESI): m/z (%): 465.3 (100) $[M+H]^+$.

N^1,N^4-Bis(3-(2,4,6-triaminopyrimidin-5-yl)propyl)terephthalimidamide hydrochloride salt (**8**): The desired compound was prepared using a procedure analogous to that described for **6** using diethyl terephthalimidate dihydrochloride (150 mg, 0.5 mmol) in EtOH (10 mL), Et_3N (0.14 mL, 1.0 mmol), and 5-(3-aminopropyl)pyrimidine- 2,4,6-triamine (205 mg, 1.1 mmol) to afford **8** as a light yellow solid (150 mg, 47 %): ^{1}H NMR (500 MH): δ =10.69 (br s, N*H*$_2$, 2H), 10.30 (br s, N*H*$_2$, 2H), 9.72 (br s, N*H*$_2$, 2H), 7.93 (s, Ar*H*, 4H), 7.25 (br s, N*H*$_2$, 8H), 7.10 (br s, N*H*$_2$, 4H), 3.13 (br s, C*H*$_2$, 4H), 1.61 (br s, C*H*$_2$, 4H), 1.17 (br s, C*H*$_2$, 4H). MS (ESI): m/z (%): 493.6 (100) $[M+H]^+$.

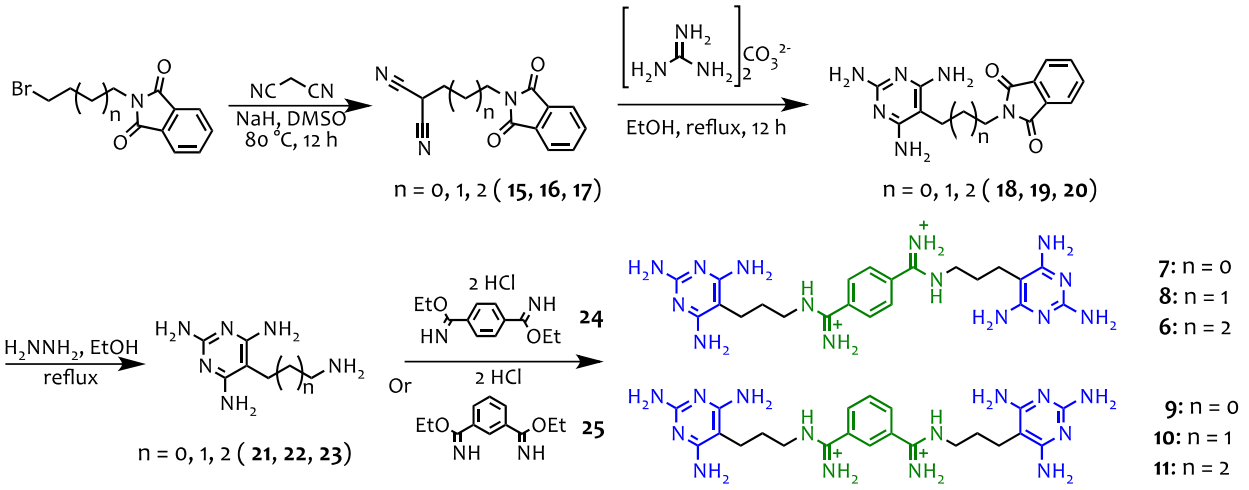
N^1,N^3-Bis(2-(2,4,6-triaminopyrimidin-5-yl)ethyl)isophthalimidamide hydrochloride salt (**9**): Et_3N (0.25 mL, 1.8 mmol) was added to diethyl isophthalimidate dihydrochloride (255 mg, 0.87 mmol) in EtOH (10 mL) cooled in an ice bath. A solution of 5-(2-aminoethyl)-pyrimidine-2,4,6-triamine (309 mg, 2.4 mmol) in EtOH (2 mL) and ethylene glycol (2 mL) was added dropwise. The reaction mixture was warmed to RT and stirred for 24 h. The reaction was monitored by TLC using 7:3 H2O/HOAc as eluent. The solvent was removed *in vacuo*. The residue was

dissolved in MeOH and loaded onto a silica gel flask column and purified by chromatography (MeOH/CH_2Cl_2 = 1:4 to 4:1). The starting material and a side product were eluted and the desired product was eluted with 4:1 MeOH/CH_2Cl_2 containing 0.04–0.08 mL of a 4M aq HCl solution in 1,4-dioxane per liter of solvent. The desired product fractions were combined, and the solvent was removed *in vacuo* to afford **9** as a HCl salt as pale yellow solid (265 mg, 50 %): mp: > 230 ºC (decomp.); ^{1}H NMR (500 MHz): δ =10.36 (br s, ArCN*H*, 2 H), 9.82 (br s, ArC=N$^+$*H*$_2$, 2H), 9.63 (br s, ArC=N$^+$*H*$_2$, 2H), 8.47 (s, Ar*H*, 1H), 8.13 (dd, J=7.8, 1.8, Ar*H*, 2H), 7.81 (t, J=7.9, Ar*H*, 1H), 7.01 (br s, CN*H*$_2$C, 8 H), 6.77 (br s, NCN*H*$_2$C, 4H), 3.48 (br t, J=7.4 Hz, NH_2C*H*$_2$, 4H), 2.74 (br t, J=7.6 Hz, CC*H*$_2$CH$_2$, 4H); MS (ESI): m/z (%): 465.3 (100) [M+H]+, 233.1 (20) $[M+H]^{2+}$.

N^1,N^3-Bis(3-(2,4,6-triaminopyrimidin-5-yl)propyl)isophthalimidamide hydrochloride salt (**10**): The desired compound was synthesized following a procedure analogous to that described for **9** using diethyl isophthalimidate dihydrochloride (150 mg, 0.5 mmol) in EtOH (10 mL), Et_3N (0.14 mL, 1.0 mmol) and 5-(3-aminopropyl)-pyrimidine-2,4,6-triamine (200 mg, 1.1 mmol) to afford **10** as a pale yellow solid (120 mg, 39%): ^{1}H NMR (500 MHz): δ =10.25 (br s, N*H*$_2$, 2H), 9.78 (br s, N*H*$_2$, 2H), 9.49 (br s, N*H*$_2$, 2H), 8.41 (s, Ar*H*, 1H), 8.06 (dd, J=6.5, 1.5, Ar*H*, 2H), 7.81 (dt, J=4, 1.5, Ar*H*, 1H), 6.40 (br s, N*H*$_2$, 8H), 6.01 (br s, N*H*$_2$, 4H), 3.48 (br t, J=6.5, C*H*$_2$, 4H), 2.38 (br t, J=8.5, C*H*$_2$, 4H), 1.67 (br q, J=3, C*H*$_2$, 4H); MS (ESI): m/z (%): 493.6 (100) $[M+H]^+$.

N^1,N^3-Bis(4-(2,4,6-triaminopyrimidin-5-yl)butyl)isophthalimidamide hydrochloride salt (**11**): The desired compound was synthesized following a procedure analogous that described for **9** using diethyl isophthalimidate dihydrochloride (150 mg, 0.5 mmol) in EtOH (10 mL), Et_3N (0.14

mL, 1.0 mmol) and 5-(4-aminobutyl)pyrimidine-2,4,6-triamine (220 mg, 1.1 mmol) to obtain **11** as a pale yellow solid (130 mg, 39%): ^{1}H NMR (500 MHz): δ =10.35 (br s, N*H*$_2$, 2H), 9.87 (br s, N*H*$_2$, 2H), 9.47 (br s, N*H*$_2$, 2H), 8.43 (s, Ar*H*, 1H), 8.09 (dd, J=7, 1.5, Ar*H*, 2H), 7.78 (t, J=7.5, Ar*H*, 1H), 7.21 (br s, N*H*$_2$, 12H), 3.49 (br s, C*H*$_2$, 4H), 2.32 (br s, C*H*$_2$, 4H), 1.78 (br s, C*H*$_2$, 4H), 1.42 (br s, C*H*$_2$, 4H); MS (ESI): m/z (%): 521.1 (100) $[M+H]^+$.



Scheme 3.2. Synthesis of **6-11**

Phthalimide 16 (n = 1): The compound was synthesized by following the procedure reported by Mascal and coworkers.1 To a suspension of 435 mg NaH (10.9 mmol) (60% in mineral oil) in 5 mL dry THF cooled in an ice bathwas added a solution of 1.2 g malononitrile (18 mmol) in 5 mL dry THF dropwise (malononitrile was recrystallized from diethyl ether). The reaction was allowed to proceed at room temperature for 30 min, and added dropwise to a suspension of N-(bromopropyl)phthalimide in 10 mL dry THF. The reaction was refluxed for 12 h. The reaction mixture was poured into 50 mL water, extracted with DCM (3 Å~ 100 mL). The combined extract was dried over Na_2SO_4 to yield 2.51 g crude containing the desired product. The crude was used for next reaction without further purification. ESI-MS (m/z): 254.1 ($[M+H]^+$, 100%.

Phthalimide 17 (n = 2): The compound was synthesized using the procedure analogous to that described for **16**. NaH (578 mg, 14.5 mmol) in dry THF (10 mL), malononitrile (1.74 g, 26 mmol) in dry THF (7 mL), N-(4- bromobutyl)phthalimide (3.71g, 13 mmol), 5.35 g crude was obtained. ESI-MS (m/z): 268.1 ($[M+H]^+$, 100% Triaminopyrimidine-phthalimide 4 (n = 0): The compound was prepared using a modification of the preparation reported by Mascal and co-workers.1 To a stirred solution of 2.5 g 2-(2-(1,3-dioxoisoindolin-2-yl)ethyl)malononitrile (10.5 mmol) in 10 mL EtOH was added 1.1 g guanidine carbonate (12.2 mmol). The reaction mixture was heated to reflux for 12 h and cooled to room temperature. The solvent was removed in vacuo. The residue was filtered, washed with a cold EtOH, and dried in vacuo to afford 1.5 g (70%) of 3 as a pale yellow solid. This compound was used for following step without further purification. ^{1}H NMR (500 MHz): δ 7.86–7.79 (m, Ar*H*, 4H), 5.42 (s, CCNH_2, 4H), 5.20 (s, NCNH_2C, 2H), 3.58 (t, J = 7.2, NCH_2, 2H), 2.57 (t, J = 7.2, CCH_2, 2H). ESI-MS (m/z): 299.1 ($[M + H]^+$, 100%). ESI-HRMS (m/z): $[M + H]^+$ calcd for $C_{14}H_{14}N_6O_2$, 299.1251, found 299.1257.

Triaminopyrimidine-phthalimide 18 (n = 0): The compound was prepared using a modification of the preparation reported by Mascal and co-workers.[1] To a stirred solution of 2.5 g 2-(2-(1,3-dioxoisoindolin-2-yl)ethyl)malononitrile (10.5 mmol) in 10 mL EtOH was added 1.1 g guanidine carbonate (12.2 mmol). The reaction mixture was heated to reflux for 12 h and cooled to room temperature. The solvent was removed *in vacuo*. The residue was filtered, washed with a cold EtOH, and dried *in vacuo* to afford 1.5 g (70%) of **3** as a pale yellow solid. This compound was used for following step without further purification. ^{1}H NMR (500 MHz): δ

7.86–7.79 (m, Ar*H*, 4H), 5.42 (s, CCNH_2, 4H), 5.20 (s, NCNH_2C, 2H), 3.58 (t, *J* = 7.2, NCH_2, 2H), 2.57 (t, *J* = 7.2, CCH_2, 2H). ESI-MS (m/z): 299.1 ($[M + H]^+$, 100%). ESI-HRMS (m/z): $[M + H]^+$ calcd for $C_{14}H_{14}N_6O_2$, 299.1251, found 299.1257.

Triaminopyrimidine-phthalimide 19 (n = 1): The compound was synthesized using the procedure analogous to that described for **18**. Crude containing 2 (2.51 g), guanidine carbonate (0.893 g, 4.9 mmol) in ethanol (20 mL). Obtain 1.3 g of 5 as a pale yellow solid. ^{1}H NMR (500 MHz): δ 7.88–7.84 (m, Ar*H*, 4H), 5.50 (s, NH_2, 4H), 5.01 (s, NH_2, 2H), 3.62 (t, J = 7.0, NCH_2CH$_2$, 2H), 2.22 (t, J = 7.5, CCH$_2$CH_2, 2H), 1.60 (q, J = 7.5, NCH$_2$CH$_2$CH_2, 2H).

Triaminopyrimidine-phthalimide 20 (n = 2): The compound was synthesized using the procedure analogous to that described for **18**. Crude containing 3 (5.35 g), guanidine carbonate (1.89 g, 10 mmol) in ethanol (30 mL). Obtain 1.22 g of compound 6 as a pale yellow solid. ^{1}H NMR (DMSO-d6): δ (ppm) 7.87–7.82 (m, Ar*H*, 4H), 5.38 (s, NH_2, 4H), 5.08 (s, NH_2, 2H), 3.57 (t, J = 7, NCH_2CH$_2$, 2H), 2.15 (t, J = 8, CCH_2CH$_2$, 2H), 1.64 (q, J = 8, NCH$_2$CH$_2$CH_2, 2H), 1.27 (t, J =7, 4.5, NCH$_2$CH$_2$CH_2, 2H).

Triaminopyrimidine 21 (n = 0): The compound was prepared using a modification of the preparation reported by Mascal and co-workers.[1] To a stirred suspension of 1.5 g of 4 (5.0 mmol) in 30 mL of EtOH was added 0.5 mL hydrazine hydrate (10.3 mmol). The reaction mixture was heated to reflux and the pale yellow solid was appeared. The reaction mixture was stirred for 12 h and cooled to room temperature. The solid was filtered, washed with 50 mL of water, and dried in vacuo to afford 0.70 g (83%) of compound 7 as a pale yellow solid. The

compound was used for next step without further purification. ^{1}H NMR (500 MHz, DMSO-d6): δ 5.47 (s, CN*H*$_2$C, 4H), 5.13 (s, NCN*H*$_2$C, 2H), 2.47 (t, J = 7.6, NH$_2$C*H*$_2$, 2H), 2.24 (t, J = 7.3, CC*H*$_2$CH2, 2H). ESI-MS (m/z): 169.1 ([M + H]+, 100%). ESI-HRMS (m/z): [M + H]+ calcd for C6H12N6, 169.1196, found 169.1203.

Triaminopyrimidine 22 (n = 1): The compound was synthesized using the procedure analogous to that described for **21**: 5 (1.3 g, 4.1 mmol), hydrazine hydrate (0.6 mL, 20.5 mmol) was refluxed in ethanol (40 mL) to obtain 800 mg of 8 as a pale yellow solid with small amount (approximate 20 % (mol/mol)) of 2,3-dihydrophthalazine-1,4-dione. ^{1}H NMR (500 MHz): δ 5.52 (s, N*H*$_2$, 4H), 5.01 (s, N*H*$_2$, 2H), 2.51 (t, J = 7.5, NC*H*$_2$CH2, 2H), 2.20 (t, J = 7, CC*H*$_2$CH2, 2H), 1.60 (q, J = 7.5, 6.5, NCH2C*H*$_2$CH2, 2H).

Triaminopyrimidine 23 (n = 2): The compound was synthesized using the procedure analogous to that described for **21**: 6 (1.2 g, 4.1 mmol), hydrazine hydrate (0.6 mL, 20.5 mmol) was refluxed in ethanol (40 mL) to obtain 635 mg of 9 as a pale yellow solid with small amount (approximate 15% (mol/mol)) of 2,3-dihydrophthalazine-1,4-dione. ^{1}H NMR (500 MHz): δ 5.45 (s, N*H*$_2$, 2H), 5.15 (s, N*H*$_2$, 2H), 2.57 (t, J = 6, NC*H*$_2$, 2H), 2.16 (t, J = 7.5, CC*H*$_2$, 2H), 1.42 (q, J = 7.5, 6, NCH$_2$C*H*$_2$, 2H), 1.30 (q, J = 7.5, 6.5, 6, NCH$_2$CH$_2$C*H*$_2$, 2H).

Diethyl terephthalimidate 24: The compound was synthesized followed the procedure reported by Wong and coworkers, the characterization data was identical to that reported.[19]

Diethyl isophthalimidate 25: To a stirred suspension of 2.9 g of 1,3-dicyanobenzene (22.6

mmol) in 35 mL absolute ethanol and 35 mL dry chloroform in ice was passed through by dry HCl gas for 5 h. The mixture was allowed warming up to room temperature and stirred for 24 h. The white solid was formed. The solid was filtered, washed with ether and dried in vacuo overnight, resulting 5.1 g (74%) of 11 as a white solid, a bisimidate esterhydrochloride. This ester was used directly in next step without further purification. ^{1}H NMR (400 MHz, CDCl3): δ 9.96 (s, Ar*H*, 1H), 9.58 (dd, J = 6.5, 1.5, Ar*H*, 2H), 8.95 (t, J = 8, Ar*H*, 1H), 5.74 (q, J = 7.5, 6, CH_2, 4H), 2.56 (t, J = 6, CH_3, 6). ESI-MS (m/z): 221.1 ([M + H]$^+$, 100%).

DM1 ligand 13: To 212 mg (0.72 mmol) of 10 in 8 mL EtOH cooled in an iced bath was added 0.25 mL Et3N (1.8 mmol). A solution of 365 mg of N2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine (1.85 mmol) in 2 mL EtOH and 2 ml ethylene glycol was added dropwise. The reaction mixture was warmed up to room temperature and stirred for 24 h. The reaction was monitored by TLC with H2O-AcOH (v/v = 7:3) as eluent. The solvent was removed in vacuo. The residue was dissolved in methanol and loaded onto a silica gel flask column and purified by chromatography (MeOH-CH2Cl2 gradient from 3:7 to 7:3). Starting material and a side product were eluted and the designed product was eluted with 7:3 MeOH-CH2Cl2 containing 0.02–0.04 mL of a 4 M aqueous HCl solution in 1,4-dioxane per liter of solvent. The desired product fractions were collected and solvent was removed in vacuo to afford 233 mg (48%) of DM1 ligand 13 as a white solid, a hydrochloride salt, m.p. > 230 °C (decomp.). 1H NMR (500 MHz): δ 10.32 (br s, ArCHNH_2, 2H), 9.89 (br s, ArCHNH_2, 2H), 9.42 (br s, ArCHN*H*, 2H), 8.42 (s, Ar*H*, 1H), 8.08 (dd, J = 6.5, 1.5, Ar*H*, 2H), 7.80 (t, J = 7.5, Ar*H*, 1H), 6.70 (s, CN*H*CH_2, 2H), 6.29 (s, CNH_2, 4H), 6.13 (s, CNH_2, 4H), 3.48 (t, J = 6, NHCH_2, 4H), 3.23 (q, J = 6.5, NHCH_2, 4H), 1.71 (q, J = 7.5, 6, 4, CH_2CH_2NH, 4H), 1.50 (q, J = 7.5, 6, 5, CH_2CH2NH, 4H). ^{13}C NMR (DMSO-d6): δ

166.6, 1662.2, 132.2, 129.7, 129.0, 43.1, 27.1, 25. ESI-MS (m/z): 523.3 ($[M+H]^+$, 40%), 262.2 ($[M+2H]^{2+}$, 70%), 175.1 ($[M+3H]^{3+,}$ 100%). ESI-HRMS (m/z): $[M + H]^+$ calcd for C22H34N16, 523.3225, found 523.3223.

3.9. Reference

(1) Raheem, O.; Olufemi, S.-E.; Bachinski, L. L.; Vihola, A.; Sirito M.; Holmlund- Hampf, J.; Haapasalo, H.; Li, Y.-P.; Udd B.; Krahe, R.; *Am. J. Pathol*. **2010**, *177*, 3025– 3036.

(2) Schoser, B.; Timchenko, L. T.; *Curr. Genomics* **2010**, *11*, 77–90.

(3) Salisbury, E.; Schoser, B.; Schneider-Gold, C.; Wang, G.-L.; Huichalaf, C.; Jin, B.; Sirito, M.; Sarkar, P.; Krahe, R.; Timchenko, N. A.; Timchenko, L. T. *Am. J. Pathol*. **2009**, 175, 748–762.

(4) Huichalaf, C.; Schoser, B.; Schneider-Gold, C.; Jin, B.; Sarkar, P.; Timchenko, L. *J. Neurosci.* **2009**, *29*, 9042– 9049.

(5) Massa, R.; Panico, M. B.; Caldarola, S.; Fusco, F. R.; Sabatelli, P.; Terracciano, C.; Botta, A.; Novelli, G.; Bernardi, G.; Loreni, F. *Neuropathol. Appl. Neurobiol.* **2010**, *36*, 275–284.

(6) Botta, A.; Caldarola, S.; Vallo, L.; Bonifazi, E.; Fruci, D.; Gullotta, F.; Massa, R.; Novelli, G.; Loreni, F.; *Biochim. Biophys. Acta Mol. Basis Dis.* **2006**, *1762*, 329– 334.

(7) Margolis, J. M.; Schoser, B. G.; Moseley, M. L.; Day, J. W.; Ranum, L. P. W. *Hum. Mol. Genet.* **2006**, *15*, 1808 –1815.

(8) Chen, W.; Wang, Y.; Abe, Y.; Cheney, L.; Udd, B.; Li, Y.-P.; *J. Mol. Biol.* **2007**, *368*, 8– 17.

(9) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. *J. Am. Chem. Soc.* **2008**, *130*, 16254– 16261.

(10) Ofori, L. O.; Hoskins, J.; Nakamori, M.; Thornton, C. A.; Miller, B. L. *Nucleic Acids Res.*

2012, *40*, 6380 –6390.

(11) Childs-Disney, J. L.; Hoskins, J.; Rzuczek, S. G.; Thornton, C. A.; Disney, M. D. *ACS Chem. Biol.* **2012**, *7*, 856 –862.

(12) Disney, M. D.; Lee, M. M.; Pushechnikov, A.; Childs-Disney, J. L. *ChemBioChem* **2010**, *11*, 375– 382.

(13) Lee, M. M.; Childs-Disney, J. L.; Pushechnikov, A.; French, J. M.; Sobczak, K.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2009**, *131*, 17464 – 17472.

(14) Gao, Z.; Cooper, T. A. *Hum. Gene Ther.* **2013**, *24*, 499– 507.

(15) Warf, M. B.; Nakamori, M.; Matthys, C. M.; Thornton, C. A.; Berglund, J. A. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18551 –18556.

(16) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16068 –16073.

(17) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. *ACS Chem. Biol.* **2013**, *8*, 1037 –1043.

(18) Jahromi, A. H.; Fu, Y.; Miller, K. A.; Nguyen, L.; Luu, L. M.; Baranger, A. M.; Zimmerman, S. C. *J. Med. Chem.* **2013**, *56*, 9471 – 9481.

(19) Wong, C.-H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez,; J. S. Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, H. Y. E.; Chan, W. Y.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 6355 – 6361.

(20) Lee, M. M.; Pushechnikov, A.; Disney, M. D. *ACS Chem. Biol.* **2009**, *4*, 345 – 355.

(21) Wong, C.-H.; Fu, Y.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Nucleic Acids Res.* **2011**, *39*, 8881 –8890.

(22) Lansiaux, A.; Tanious, F.; Mishal, Z.; Dassonneville, L.; Kumar, A.; Stephens, C. E.; Hu,

Q.; Wilson, W. D.; Boykin, D. W.; Bailly, C. *Cancer Res.* **2002**, *62*, 7219 –7229.

(23) Hung, M.; Patel, P.; Davis, S.; Green, S. R.; *J. Virol.* **1998**, *72*, 4819– 4824.

(24) Marcheschi, R. J.; Tonelli, M.; Kumar, A.; Butcher, S. E. *ACS Chem. Biol.* **2011**, *6*, 857 – 864.

(25) Warf, M. B.; Berglund, J. A. *RNA* **2007**, *13*, 2238 –2251.

(26) Childs-Disney, J. L.; Yildirim, I.; Park, H.; Lohman, J. R.; Guan, L.; Tran, T.; Sarkar, P.; Schatz, G. C.; Disney, M. D. *ACS Chem. Biol.* **2014**, *9*, 538 –550.

(27) Mascal, M.; Hansen, J.; Fallon, P. S.; Blake, A. J.; Heywood, B. R.; Moore, M. H.; Turkenburg, J. P. *Chem. Eur. J.* **1999**, *5*, 381 –384.

(28) Vichai, V.; Kirtikara, K. *Nat. Protoc.* **2006**, *1*, 1112 – 1116.

(29) Yuan, Y.; Compton, S. A.; Sobczak, K.; Stenberg, M. G.; Thornton, C. A.; Griffith, J. D.; Swanson, M. S. *Nucleic Acids Res.* **2007**, *35*, 5474 –5486.

Chapter 4

Multi-targeting small molecules of Myotonic Dystrophy type 1[1]

4.1. Introduction

Drug discovery efforts traditionally place a high premium on agents that operate on a single target with high selectivity and affinity. However, rapid advances in "omics" have revealed the complexity of many diseases, especially cancer, where as many as 500 gene products may be dysregulated.[1] In such cases, a "magic bullet" approach may be quite limited. Indeed, studies have suggested that in at least some cases less selective drugs exerting their pharmacologic effect on multiple targets can be superior to those with narrow activity profiles.[2] These realizations have led some to suggest a paradigm shift from "single drug, single target" to polypharmacologic or multi-target drug discovery (MTDD) approaches.[2,3] The main challenge in MTDD is the need to design a drug that modulates multiple disease targets simultaneously. Especially difficult is creating fused, hybrid structures whose different molecular segments recognize different targets. A much simpler strategy is to tether together two or more structural domains with different biological activities.[4,5] These multiple ligands may be conjugates of two known inhibitors (e.g., of two signaling pathways) or a binding unit and a chemically reactive group.

Although MTDD efforts have largely focused on protein targets, the increasing importance of RNA as a therapeutic target makes it an excellent candidate for MTDD. We were particularly attracted to myotonic dystrophy type 1 (DM1) because its complicated disease pathogenesis is increasingly well-understood, providing well-defined DNA, RNA, and protein targets for a small-molecule MTDD approach. DM1 is an incurable, multisystemic neuromuscular disease

[1] The material in this chapter was adapted from the following publication:
Nguyen, L.; Luu, M. L.; Peng, S.; Serrano, F. J; Chan, W.-Y.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2015**, *137*, 14180-14189.

that is caused by an abnormal expansion of the CTG trinucleotide repeats (CTG^{exp}) in the 3′-untranslated region of the *DMPK* gene on chromosome 19q13 (see Figure 4.1a, top box).[6] This expanded DNA, which can reach 50−2000 CTG repeats, yields an expanded CUG RNA transcript (CUG^{exp}) that sequesters the alternative-splicing regulator muscleblind-like protein (MBNL), leading to splicing defects and disease symptoms. We and others have developed small molecules that inhibit the MBNL1 sequestration by CUG^{exp}.[7-14] However, a recently expanded view of DM1 pathogenesis has suggested that additional CUG^{exp}-induced toxic pathways must be considered for the disease phenotype to be fully reversed. In particular, the CTG·CAG repeats undergo bidirectional transcription, producing two transcripts, CUG^{exp} and CAG^{exp},[15,16] that both undergo repeat-associated non-ATG (RAN) translation, generating multiple toxic homopeptides.[17,18] Further, it was shown that CUG^{exp} disrupting the translation of the MEF2

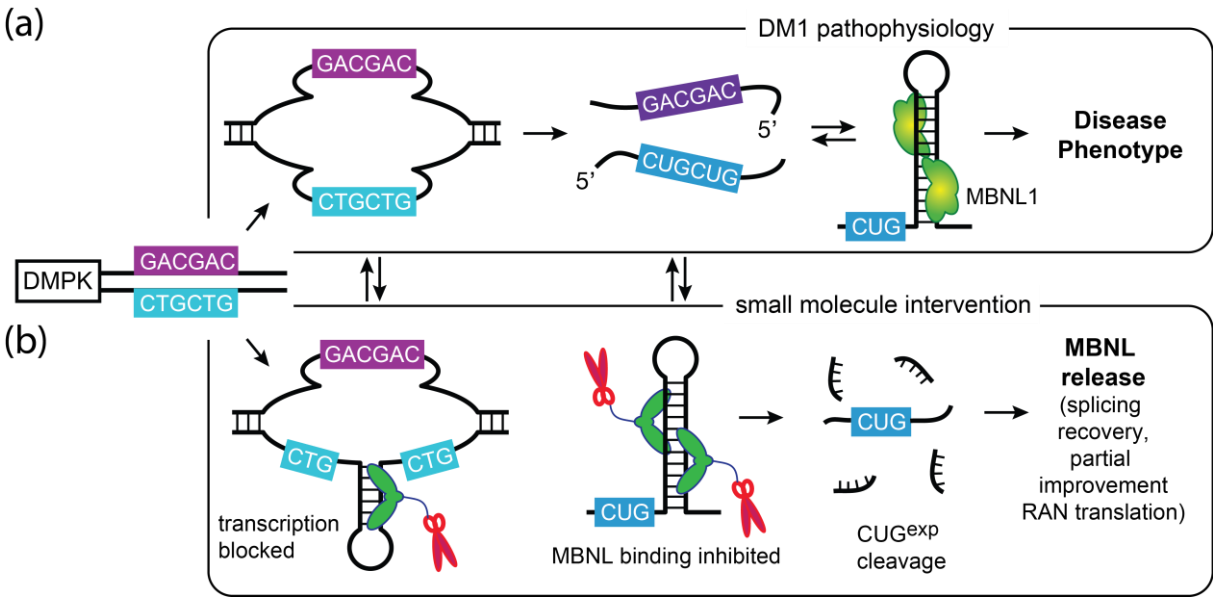
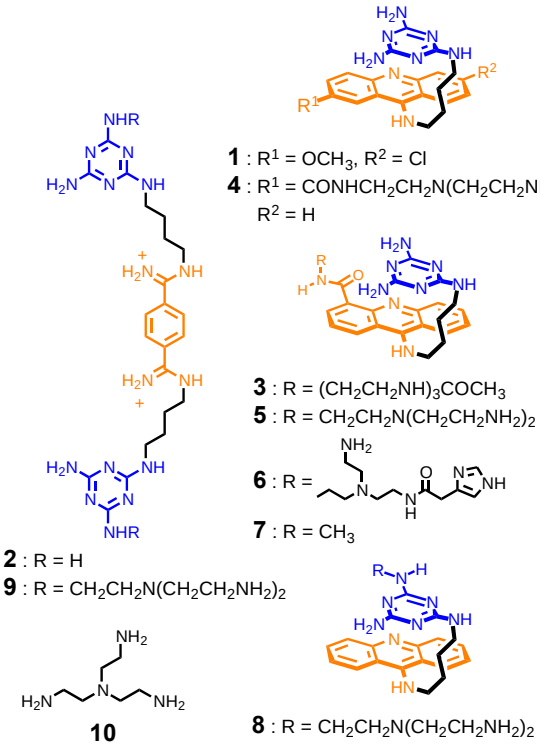


Figure 4.1. Schematic illustration of DM1 pathobiology and multi-target treatment. (a) RNA gain-of-function disease pathogenesis. The expanded DNA trinucleotide repeat (CTG^{exp}) undergoes transcription to form a CUG^{exp} hairpin that sequesters MBNL proteins (e.g., MBNL1). The MBNL level depletion causes splicing defects of more than 100 pre-mRNAs, resulting in disease symptoms. RAN translation of CUG^{exp} and CAG^{exp} generates toxic homopeptides. (b) Small-molecule intervention. Small molecules (green) target the CTG^{exp} hairpin, inhibiting production of CUG^{exp}. Any CUG^{exp} formed is bound by small molecules, inhibiting MBNL and other protein sequestration. Cleaving functionality (red) processes the toxic RNA. All three small-molecule interventions free MBNL for its normal function. Likewise, other toxic pathways induced by the CUG^{exp} are eliminated.

protein, which affects multiple levels of mRNA and microRNA in human DM1 heart tissues.[19] Furthermore, the discovery of other proteins involved in the formation of MBNL1−CUG^{exp} foci suggests that other toxic pathways may be induced by CUG^{exp}.[20,21]

The studies above suggest that a multi-target drug approach, especially one that degrades the toxic CUG^{exp} or inhibits its formation, may be more effective. Agents that perform one of these two functions are known. Thus, Cooper[22] and Thornton[23] reported antisense agents that induce CUG^{exp} cleavage via an RNase H-mediated mechanism and are currently in clinical trials, and Disney developed a small molecule that photodegrades the toxic transcript.[24] Analogues of pentamidine were reported by Berglund to inhibit the synthesis of CUG^{exp} by binding to the $(CTG{\cdot}CAG)_n$ duplex.[25] We report herein a rational MTDD effort leading to small molecules that intervene in three separate steps in the DM1 pathobiology, suppressing CUG^{exp} mRNA levels and reversing the disease phenotype in DM1 model cells and a DM1 Drosophila model.

Chart 4.1.



4.2. Design and synthesis of multi-target small molecules

The overall approach was to develop agents able to intervene in the DM1 pathogenic mechanism in three ways: (1) targeting CTG^{exp} to inhibit the transcription to CUG^{exp}, (2) targeting CUG^{exp} to inhibit the MBNL sequestration, and (3) hydrolytically degrading the CUG^{exp} with RNase-like catalytic functionality (Figure 4.1). Our group previously reported two classes of rationally designed agents (e.g. **1** and **2**) that selectively bound CUG^{exp} and inhibited its interaction with MBNL with low micromolar K_i values.[11,12,26] Both ligands feature triaminotriazine moieties to recognize the UU mismatch in duplex CUG^{exp} and provide sequence selectivity, whereas the acridine group in **1** and bisamidinium unit in **2** were selected to drive the association by A-form RNA intercalation and groove binding, respectively. Although **1** was not cell-permeable and poorly water-soluble, analogue **3** utilized the polyamine transport system to enter cells.[27,28] The bisamidinium unit in **2** was chosen partly because it was reported to localize in cell nuclei. Both **2** and **3** dissolved the MBNL−CUG^{exp} nuclear foci and partially rescued splicing defects of *IR* and *cTNT* minigenes in DM1 model cell cultures (Chart 4.1).[12,27] In comparison to **1**, bisamidinium-containing **2** showed similar *in vitro* inhibition potency, but with lower toxicity, higher water solubility, and better cell uptake. Nonetheless, the acridine ligands are inherently multi-targeting because of their ability to bind both the DNA and RNA causing DM1. Thus, agent **1** complexes an oligonucleotide 10-mer containing a single dCTG and rCUG site with the binding constant K_D = 0.39 and 0.43 nM, respectively.[11] In contrast, agent **2** is selective for CUG^{exp}, showing no detectable binding of dCTG by isothermal titration calorimetry (ITC) (*vide infra*). Beyond their ability to inhibit the sequestration of MBNL and other proteins, both **1** and **2** have the potential to become CUG^{exp} cleaving agents with attachment of suitable functionality.

Numerous small-molecule mimics of RNase A were developed over the past few decades.[29,30]

Many use the active-site functional groups found in RNase A in an effort to mimic its well-established acid−base mechanism of action.[31] Although none of the mimics cleaved RNA as effectively as RNase A, those containing at least one ammonium ion and an imidazole or amino group were the most promising. Based on this information, we designed agents **4**−**6**, **8**, and **9** with side chains at the 2- or 4-position of the acridine ring (**4**−**6**) or attached to the triaminotriazine rings (**8** and **9**).

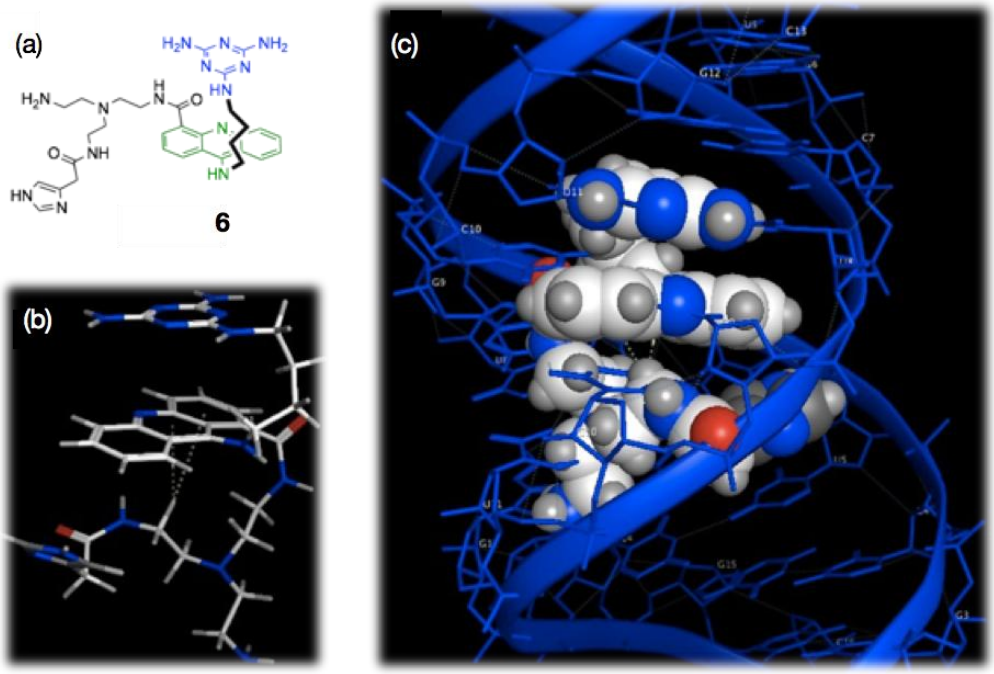


Figure 4.2. Representative modeling of CUG^{exp} cleaving agent **6** with $(CUG)_6$ using MOE**.** All synthesized CUG^{exp} cleaving agents were modeled on $(CUG)_6$ (PDB: 3GM7) and $(CUG)_{15}$ (for ligand **9**) showing that the cleaving units can reach the phosphate backbone. (a) Structure of ligand **6.** (b) Minimized energy ligand **6** in water. (c) Modeling the binding of **6** on $(CUG)_6$. The bound conformer of ligand 6 was a local minimum with the potential energy of 131 kcal/mol. Computational work is intended only to show that these flexible compounds are able to reach the backbone.

The ability of these potentially catalytic functional groups to perform in the desired fashion was examined by modeling. Thus, the ligands were docked to binding sites prepared from the published X-ray analysis of $(CUG)_6$ (PDB: 3GM7) using MOE (Figure 4.2), and each was found to reach at least one putative scissile phosphate bond. In collaboration with Long Luu, cleaving agents **4**−**6**, **8**, and **9** were synthesized following generally well-established procedures (Schemes

4.1−4.4, Materials and Methods). The synthesis broadly involved conjugation of di-Boc-protected tris(2-aminoethyl)amine with the appropriate acridine acid chloride or chlorotriazine. The imidazole-bearing agent **6** was prepared from **5** by coupling one amino group with imidazole-4-acetic acid. The additional functional groups improved the water solubility, cell penetration, and affinity toward CUG^{exp} (Figure 4.3).

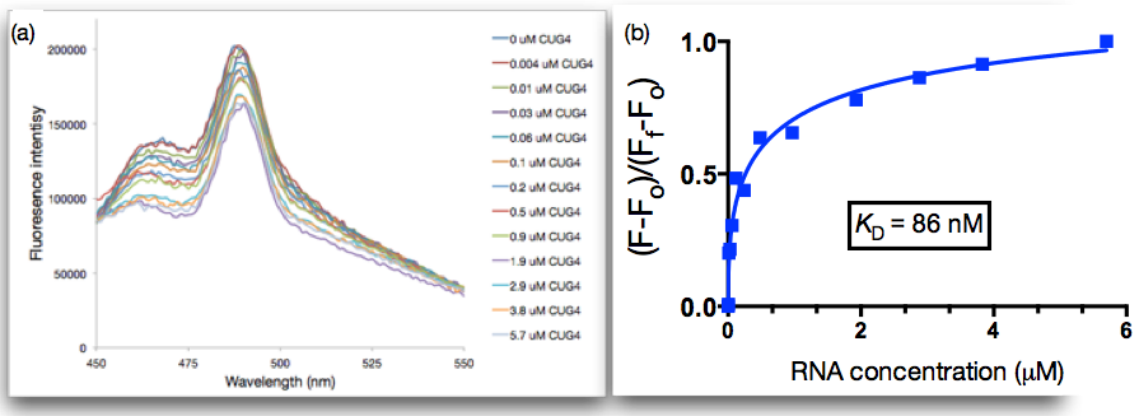


Figure 4.3. Binding study of ligand **5** and $(CUG)_4$ by reverse-titrations. **(a)** Change in fluorescence intensity of **5** with increasing $(CUG)_4$ concentration. **(b)** The plot of fluorescence change versus the concentration of $(CUG)_4$ in a range from 0-6 μM provided an apparent dissociation constant K_D of 86 nM as an average of two independent experiments. The sample was excited at 420 nm and the maximum emission was monitored at 490 nm. Experiments were performed following a reported protocol.[1]

4.3. DNA target activity: *In vitro* transcription inhibition of $(CTG{\cdot}CAG)_n$ by DM1 multi-target small molecules

As described above, simple ligand **1** binds CTG sites tightly, whereas **2** does not. Nonetheless, the additional ammonium groups in **9** might potentially increase its affinity for both CTG DNA and CUG RNA, so we examined its potential for inhibition of CTG^{exp} transcription along with that of ligands **5** and **6**. The *in vitro* $(CTG{\cdot}CAG)_{74}$ transcription assays utilized a T7 promoter located in the upstream region of the repeats.[32] Ligands **5** and **9** strongly inhibited the production of CUG^{exp} and in a dose dependent manner, whereas ligand **6** showed less inhibition (Figure

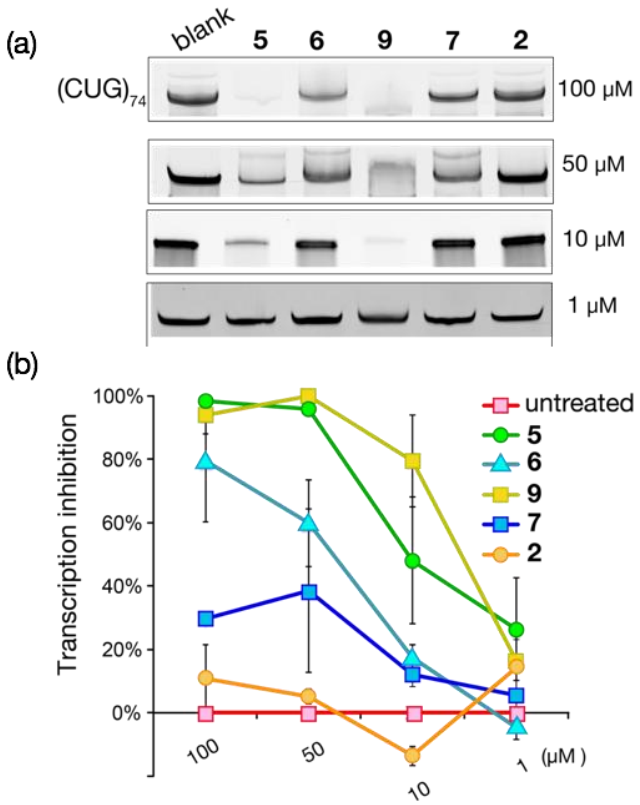


Figure 4.4. *In vitro* transcription of (CTG·CAG)74. Ligands at different concentrations were incubated with 15 ng of linearized plasmids in T7 RNA polymerase mixture at 37 °C. After 2 h, the reaction mixture was loaded on a 8% denaturing gel. (a) Transcription gels of ligands at different concentrations. (b) Plot of the percentage of transcription inhibition versus ligand concentrations. The error bars represent standard errors of mean of three independent experiments.

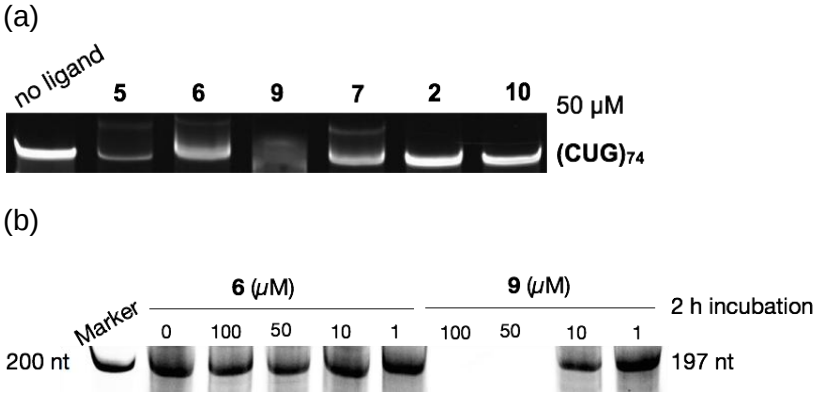


Figure 4.5. *In vitro* transcription gels. (a) Control ligand **10** showed negligible effect on *in vitro* transcription of $(CTG·CAG)_{74}$. (b) Effects of ligands **6** and **9** on the transcription of control plasmids that contain no CTG·CAG repeats. Ligand **6** showed negligible transcription inhibition whereas **9** strongly inhibited the formation of control plasmids at 100 and 50 μM. However, the transcription inhibition of **9** is significantly reduced to less than 30% at 10 μM.

4.4). Control ligands **2**, **7**, and **10** had negligible effect on the transcription of $(CTG·CAG)_{74}$

(Figure 4.4 and Figure 4.5a). To test the selectivity and potentially the target of the inhibition, we performed similar experiments with two separate control plasmids each lacking repeats (see Materials and Methods for details). Ligand **6** showed negligible transcription inhibition at all tested concentrations, which ranged from 1 to 100 μM, whereas **9** strongly inhibited the

Table 4.1. Equilibrium Dissociation Constants (apparent K_D , μM) of Ligands **2** and **9** to Various Oligonucleotides Determined by ITC[a]

Hairpin/Duplex	**2**	**9**
$(CUG)_{12}$	8 ± 2[b]	6 ± 4
$d(CTG)_{12}$	nb	5 ± 1
$d(CAG)_{12}$	nb	nb
$d(CTG{\cdot}CAG)_{12}$	nb	nb

[a]Apparent K_D values were determined from at least three independent experiments; nb indicates no detectable binding. [b]Data from Wong (ref 12).

transcription of the control plasmids at higher concentrations (i.e., 50 and 100 μM) (Figure 4.5b). However, **9** did show some selectivity with only 30% inhibition of the control plasmids observed at 10 μM, whereas ca. 80% inhibition was obtained for the plasmid containing $(CTG{\cdot}CAG)_{74}$ (Figure 4.4b). In collaboration with Julio Serrano, we determined the binding affinity of ligands **2** and **9** toward the DM1 RNA and DNA. The lack of binding affinity shown by **2** toward

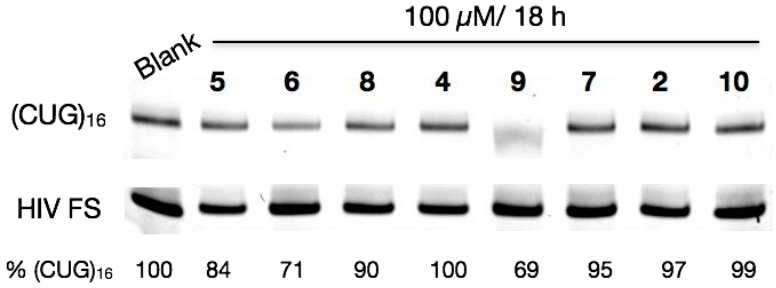


Figure 4.6. Cleavage screening gel. $(CUG)_{16}$ (100 nM) was incubated with ligands (100 μM) in a Tris buffer (pH 7.4) containing 150 mM NaCl and 2 mM $MgCl_2$, 18 h, 37 °C. HIV FS RNA (1 pmol) was used as an internal standard. The reaction mixture was loaded onto a 20% denaturing PAGE gel, and RNA was detected by poststaining with EtBr. Normalized % $(CUG)_{16}$ intensity was reported as the average value of two independent experiments.

$d(CTG)_{12}$, $d(CAG)_{12}$, and $d(CTG{\cdot}CAG)_{12}$ (Table 1), inhibits the transcription of CTG^{exp} by stabilizing its hairpin structure. Indeed, no strong binding was detected for $d(CAG)_{12}$ or $d(CTG{\cdot}CAG)_{12}$. The *in vitro* transcription experiments using $(CTG{\cdot}CAG)_{74}$ and compounds **2** , **5**–**7** , and **9** were repeated but with $(CUG)_{16}$ added as a competitor. In each case, the transcription inhibition of ligands was reduced, particularly for **6**.

4.4. RNA target activity: CUG cleavage by DM1 multi-target small molecules

The potential cleavage activity of **4**–**9** was screened using a simple gel shift assay with $(CUG)_{16}$. Thus, each agent at a final concentration of 100 μM was incubated for 18 h with unlabeled $(CUG)_{16}$ at pH 7.4. The mixture was separated on an RNA denaturing gel and stained with EtBr (Figure 4.6). No loss of $(CUG)_{16}$ intensity was observed upon treatment with **2**, **7**, or tris(2-

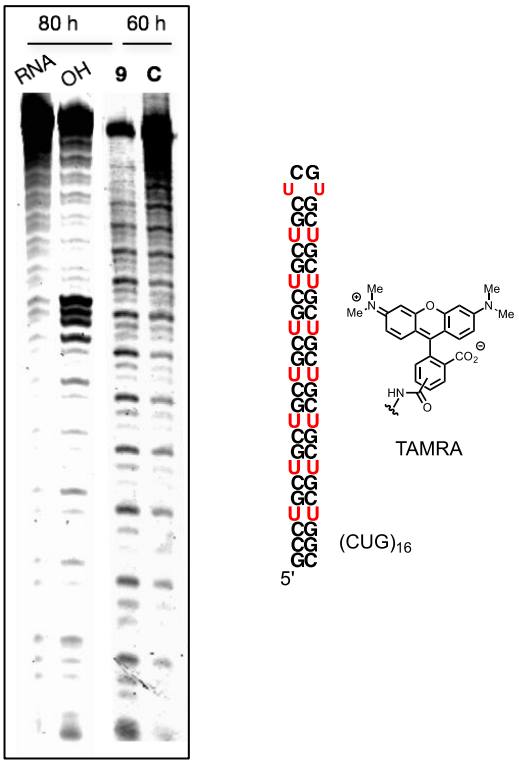


Figure 4.7. TAMRA−$(CUG)_{16}$ cleavage gel, $(CUG)_{16}$ and TAMRA structures . Ligand **9** or a control **C** as mixture of **2** (100 μM) and **10** (100 μM) was incubated with T-$(CUG)_{16}$ (100 nM) in a Tris buffer (50 mM, pH 7.4) supplemented with 150 mM NaCl and 2 mM $MgCl_2$. OH is a control with RNA incubated in a buffer at pH 10.6. The reaction mixture was run on a 20% denaturing PAGE gel.

aminoethyl)amine (**10**), demonstrating that a polyamine or the CUG-targeting acridine or bisamidinium ligand on its own is insufficient to alter the $(CUG)_{16}$. In contrast, a decrease in $(CUG)_{16}$ intensity was observed for those ligands, i.e., **5**, **6**, **8**, and **9**, that contain a least one amino group, consistent with RNA cleavage (Figure 4.6). At least qualitatively, agents **6** and **9** appeared to be most active.

To observe potential cleavage fragments that were not observable in the screening gel because of the relative insensitivity of the EtBr post-staining, we performed similar experiments using 5′ -TAMRA-labeled $(CUG)_{16}$ (T-$(CUG)_{16}$). As seen in Figure 4.7, after 60 h of incubation, agent **9** shows a large number of bands. Interestingly, the intensity pattern corresponds to the repeat sequence, with every fourth band significantly more intense. This pattern may indicate a specific positioning of the catalytic groups by the ligand or possibly a higher reactivity of the UU mismatch site. Similar, although less distinct, patterns were seen at shorter (20 h) incubation times. RNA fragments were observed for a control incubation using a combination of ligand **2** and tetraamine **10**, but the reaction was much slower. Given the simplicity and ease of quantifying the loss of unlabeled $(CUG)_{16}$ in the screening assay above, agent **9** was re-examined

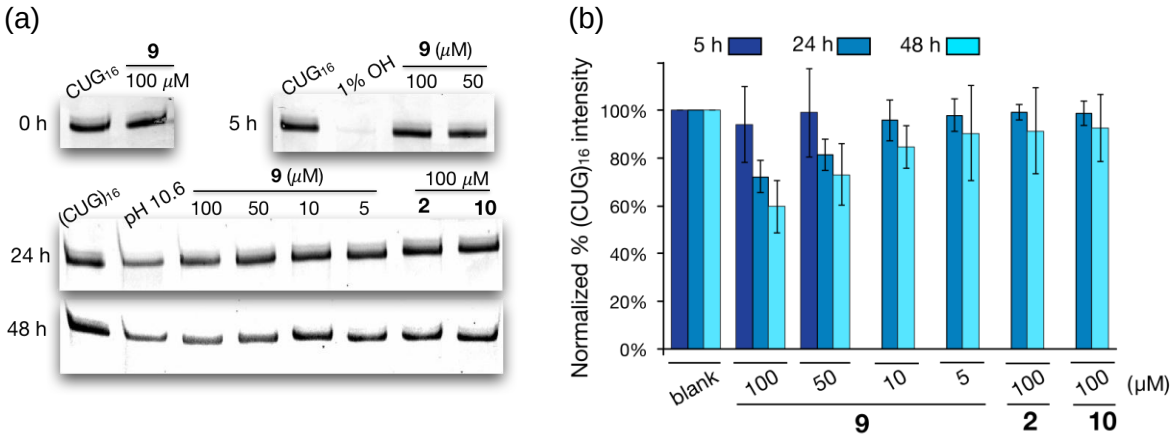


Figure 4.8. Time- and dose-dependent cleavage experiments of ligand **9** using $(CUG)_{16}$. (a) EtBr post-staining gels of time- and dose-dependent cleavage experiments. (b) Quantitative analysis of changes in $(CUG)_{16}$ intensity. The reaction mixture was loaded on a 20% denaturing PAGE gel. The percentage of $(CUG)_{16}$ intensity was normalized to the control (no compound). Error bars represent standard errors of mean of three independent experiments.

at four different concentrations (5 − 100 μM). As seen in Figure 4.8, loss of the $(CUG)_{16}$ band was both time- and dose-dependent.

To better determine the origin of the loss of CUG^{exp} mRNA, the reaction of $(CUG)_4$ with a 3′ -TEG-biotin tag and **6** was monitored by MALDI. After a 5 h incubation with **6**, fragment peaks were observed with lower m/z values, but no major change was seen in control samples

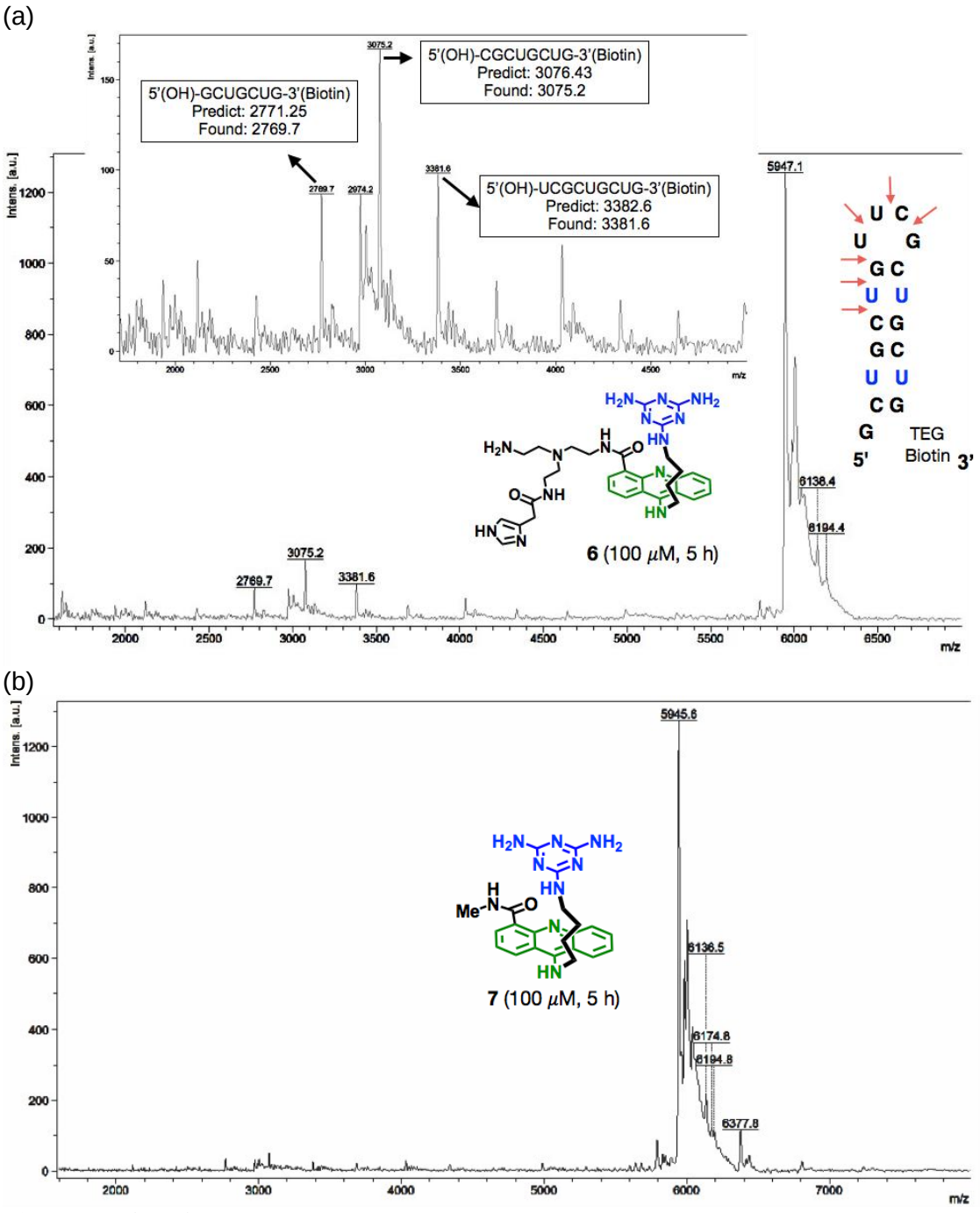


Figure 4.9. (cont.)

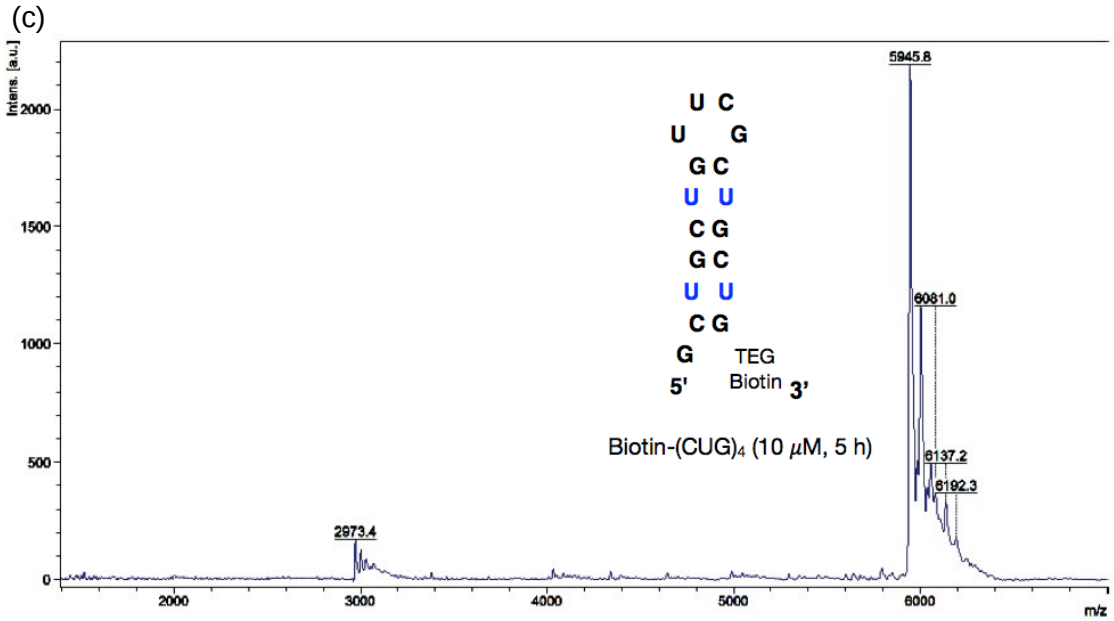


Figure 4.9. MALDI spectra of $(CUG)_4$ cleavage of ligand **6.** The cleavage sites corresponding to the most abundant peaks are indicated with red arrows. (a) RNA incubated with **6**. (b) RNA incubated with control ligand **7**. (**c**) RNA only. Biotin-labeled $(CUG)_4$ (10 μM) was incubated with ligands (100 μM) or water in cleavage buffer (50 mM Tris buffer (pH 7.4), 150 mM NaCl, 2 mM $MgCl_2$) for 5 h at 37 °C. The reaction mixture was cleaned up using C18-Ziptip and subjected to MALDI.[2] The fragments from MALDI-TOF were compared with the predicted values. The values of m/z with a mass-matching tolerance of 350 ppm were considered identical.

(Figure 4.9). The m/z values found matched the calculated m/z corresponding to hydrolysis products with 3′-hydroxyl end groups and selective cleavage in the loop and immediately adjacent to the loop.

Cleavage experiments at different pH values and different concentrations of Mg^{2+} were carried out. The data from these experiments showed that the cleavage activity of ligand **9** was Mg^{2+} -independent and increased with increasing pH (Figure 4.10). The higher pH increases the concentration of the anticipated active species containing one amino group and one ammonium ion; however, the higher pH also increases the background hydroxide-catalyzed reaction.

4.5. Selectivity of CUG^{exp} cleaving agents

Off–target activity is a concern for any therapeutic agent but especially one designed to chemically alter its target (e.g., cleave RNA). To test the selectivity of agents **6** and **9**, the

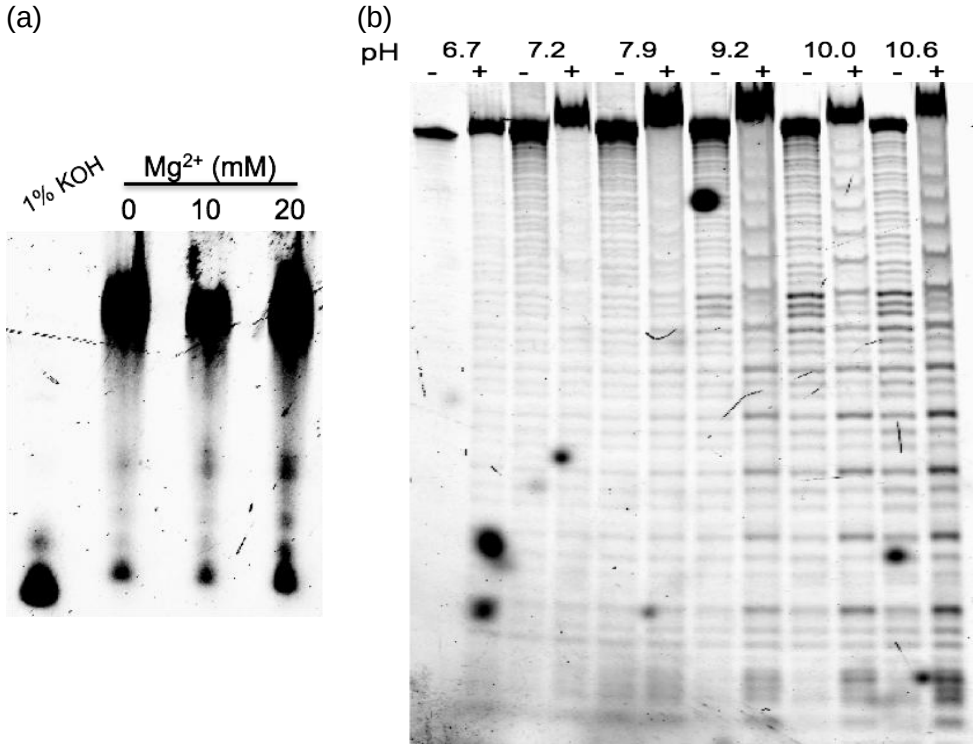


Figure 4.10. Factors affecting RNA cleavage activity of ligand **9**. $(CUG)_{16}$ cleavage study of **9** at difference concentrations of Mg^{2+}(**a**) and pH (b). T-$(CUG)_{16}$ was incubated with **9** (100 μM) for 5 h in buffers containing 0, 10, and 20 mM Mg^{2+} or at different pH values in a range from 6.7-10.6. The reaction was quenched by 8 M urea and the samples were loaded on a 20% denaturing PAGE gel. (-) no ligand added, (+) ligand with the concentration of 100 μM added.

cleaving gel assay described above was applied to other RNA targets, specifically cTNT32, $(CCUG)_8$, and HIV-1 frameshift site RNA (HIV FS) (Figure 4.11a). Both **6** and **9** were quite selective. Thus, neither agent showed detectable cleavage of cTNT32 or HIV FS, despite both of these RNAs having structures loosely analogous to CUG^{exp} (Figure 4.11). Both similarly adopt stem-loop structures with internal loops within the stem, although obviously of different sequence. Differences between **6** and **9** were observed with the former cleaving $(CCUG)_8$, the toxic RNA involved in myotonic dystrophy type 2 (DM2) (Figure 4.11). In contrast, **9** showed no

activity toward $(CCUG)_8$. The ability to process the RNA directly parallels the corresponding ligand−RNA binding affinities. Thus, **1**, which contains the acridine−triazine core of agent **6** , complexes $(CCUG)_6$[33] and $(CUG)_n$ sequences (see also Figure 4.3). Likewise, **2** (the core of agent **9**) complexes $(CUG)_n$ but showed no affinity toward $(CCUG)_8$, cTNT32, or HIV FS.[12] These data are consistent with the catalytic functionality being brought into proximity of the RNA through selective binding.

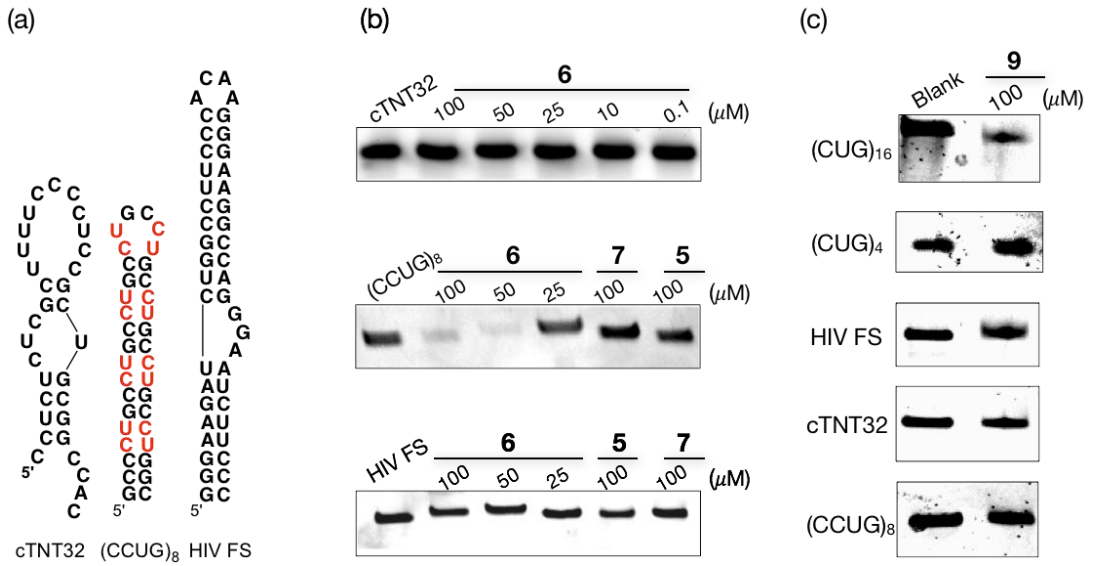


Figure 4.11. Selectivity study. (a) Structures of RNA oligonucleotides tested in the selectivity study. (b, c) Gels of potential cleavage of alternative RNA targets by ligands **6** and **9**. RNA (100 nM) was incubated with ligands for 19 h at 37 °C. The reaction mixture was loaded on a 20% denaturing PAGE gel, followed by post-staining with EtBr.

4.6. Bioactivity of multi-target agents in DM1 model cells

The ability of **5**, **6**, and **9** to disrupt the MBNL1−CUG^{exp} interaction in cells was evaluated using model DM1 cells. Thus, HeLa cells were transfected with GFP-DT0 or GFP-DT960 plasmids that contain 0 or 960 interrupted CTG repeats, respectively, in exon 15 of a truncated *DMPK* gene.[19,22] The plasmids express both GFP protein and CUG^{exp} under the activation of doxycycline (Dox), with GFP being used as a marker for successful transfection and expression

of the plasmids in the cells. Treatment of the DM1 model cells with ligands at 50 μM for 48 h was followed by analysis using confocal microscopy. As seen in the representative images in Figure 4.12, each of the three agents inhibited nuclear foci formation, leading to the dispersion of MBNL1 protein throughout the nucleus.

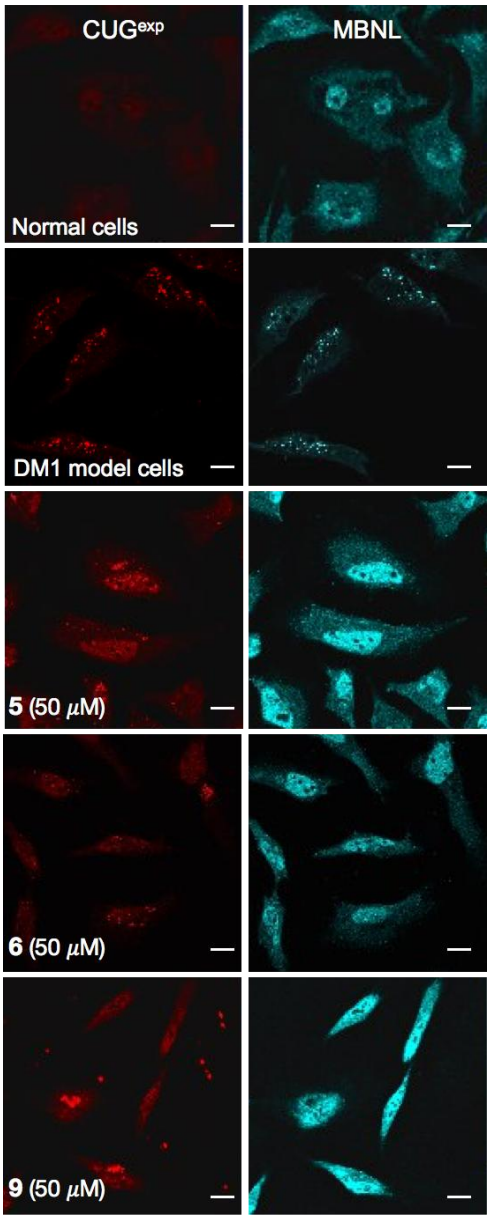


Figure 4.12. Foci dispersion by agents **5**, **6**, and **9**. DM1 model cells were incubated with 50 μM ligand for 48 h. Cells were fixed, and CUG^{exp} was stained with Cy3-$(CAG)_{10}$, and MBNL1 was probed with mouse anti-MBNL followed by staining with goat anti-mouse Alexa 647 secondary antibody. The scale bar is 10 μm.

Because **9** is less toxic toward HeLa cells than **5** and **6**, it was selected for a splicing study of insulin receptor (*IR*) minigene. Splicing of *IR* minigene is misregulated with the abnormal

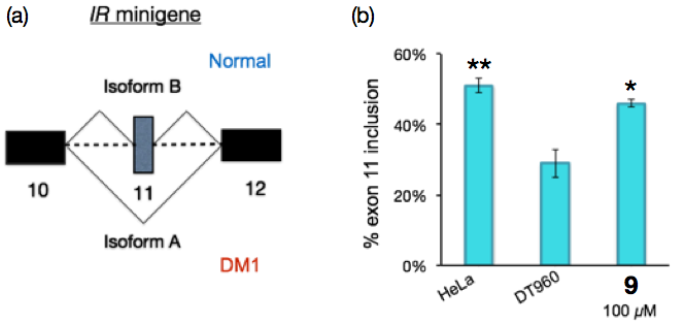


Figure 4.13. Ligand **9** rescued splicing defect of *IR* minigene. (a) Mis-splicing of *IR* minigene in DM1 model cells. (b) Percentage of exon 11 inclusion in HeLa (healthy model cell). DT960 (HeLa co-transfected with *IR* plasmid and DT960 plasmid expressing $(CUG)_{960}$), and DT960 cells treated with **9** for 3 d. The splicing experiments were performed following a reported protocol.[3] Error bars represent the standard error of mean from at least three independent experiments. * $P < 0.05$, ** $P < 0.005$ (two tailed *t*-test).

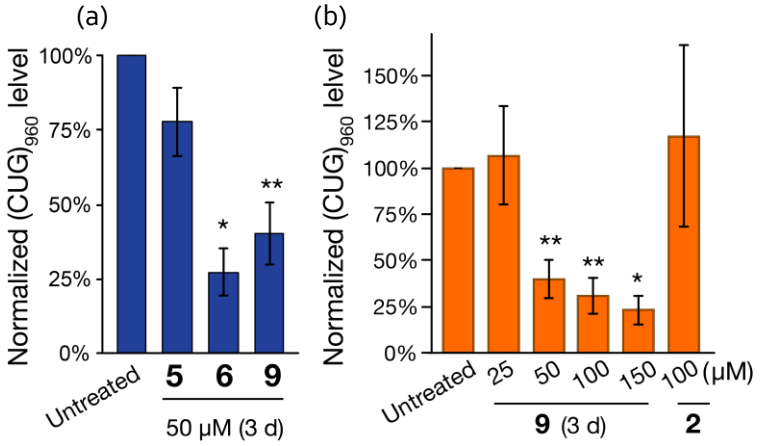


Figure 4.14. CUGexp mRNA levels in DM1 model cells. (a) Effects of ligands 5, 6, and 9 on CUGexp mRNA levels at 50 μM for 3 day treatment. (b) Ligand 9 reduced CUGexp mRNA levels in a dose-dependent fashion. Error bars represent the standard error of mean of at least three independent experiments; *$P < 0.01$, **$P < 0.005$ (two tailed *t*-test).

exclusion of exon 11 because of the MBNL1 depletion in DM1 cells.[34] DM1 model cells, in this case HeLa cells co-transfected with plasmids containing *IR* minigene and $(CTG)_{960}$, were treated with **9** at 100 μM for 3 days, leading to a 77% rescue of the IR splicing defect (Figure 4.13). Encouraged by the promising results from the *in vitro* cleavage and the ability of **5**, **6**, and **9** to

enter cells and dissolve MBNL1− CUG nuclear foci, we performed experiments to study whether the agents could control the level of toxic CUG^{exp} in cells using a reported protocol.[19,22] Two sets of cells, transfected with either GFP-DT0 or GFP-DT960 plasmids, were incubated with **5**, **6**, or **9** for 3 d. The CUG^{exp} mRNA level was determined by measuring the level of exon 15 upstream of the CUG^{exp} relative to PABP mRNA as a control, followed by normalizing the

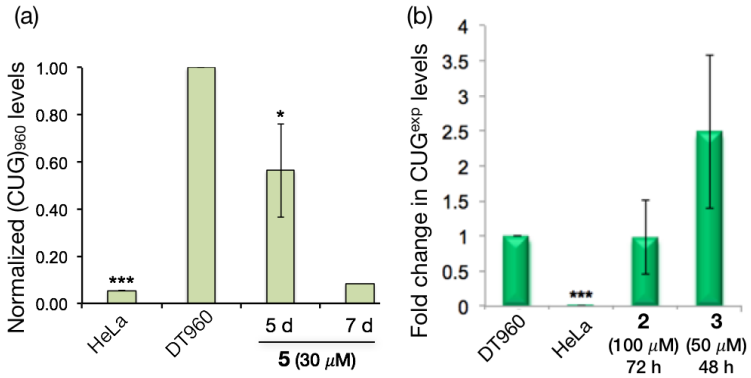


Figure 4.15. CUG^{exp} mRNA level study. (a) CUG^{exp} mRNA levels after treatment with ligand **5**. (b) Fold change in CUG^{exp} expression levels in DM1 model cells treated with control ligands **2** and **3**. CUG^{exp} levels from untreated DM1 model cells as HeLa cells transfected with plasmid containing $(CTG)_{960}$ was normalized to 1. Ligand **2** at 100 μM showed no effect on the level of $(CUG)^{exp}$ after 72 h incubation, whereas treatment of DM1 model cells with **3** at 50 μM for 48 h resulted in an increase in CUG^{exp} mRNA level. The error bars represent standard deviation of three independent triplicates. * $P < 0.05$, *** $P < 0.001$ (two tailed *t*-test).

values to the levels measured from untreated cells. As seen in Figure 4.14a, there was a 60−70% reduction in CUG^{exp} levels in cells treated with **6** and **9** at 50 μM. Ligand **5** showed only a negligible change of CUG^{exp} levels, although a longer incubation time of 5 and 7 d led to a significant decrease in the level of toxic CUG^{exp} RNA (Figure 4.15a). Because of its lower toxicity, the dose dependence of **9** was examined by treating cells with four different concentrations ranging from 25 to 150 μM. As seen in Figure 14.4b, **9** clearly regulates the cellular CUG^{exp} mRNA levels in a dose-dependent manner. The control ligands **2** and **3** did not show a significant effect on the level of toxic RNA (Figure 4.14b and Figure 4.15b).

4.7. Effects of ligand 9 on the disease phenotypes in DM1 *Drosophila*

The data presented above indicate that **9** is able to engage each of three small-molecule intervention pathways outlined in Figure 4.1b, and it was found further to be relatively nontoxic.

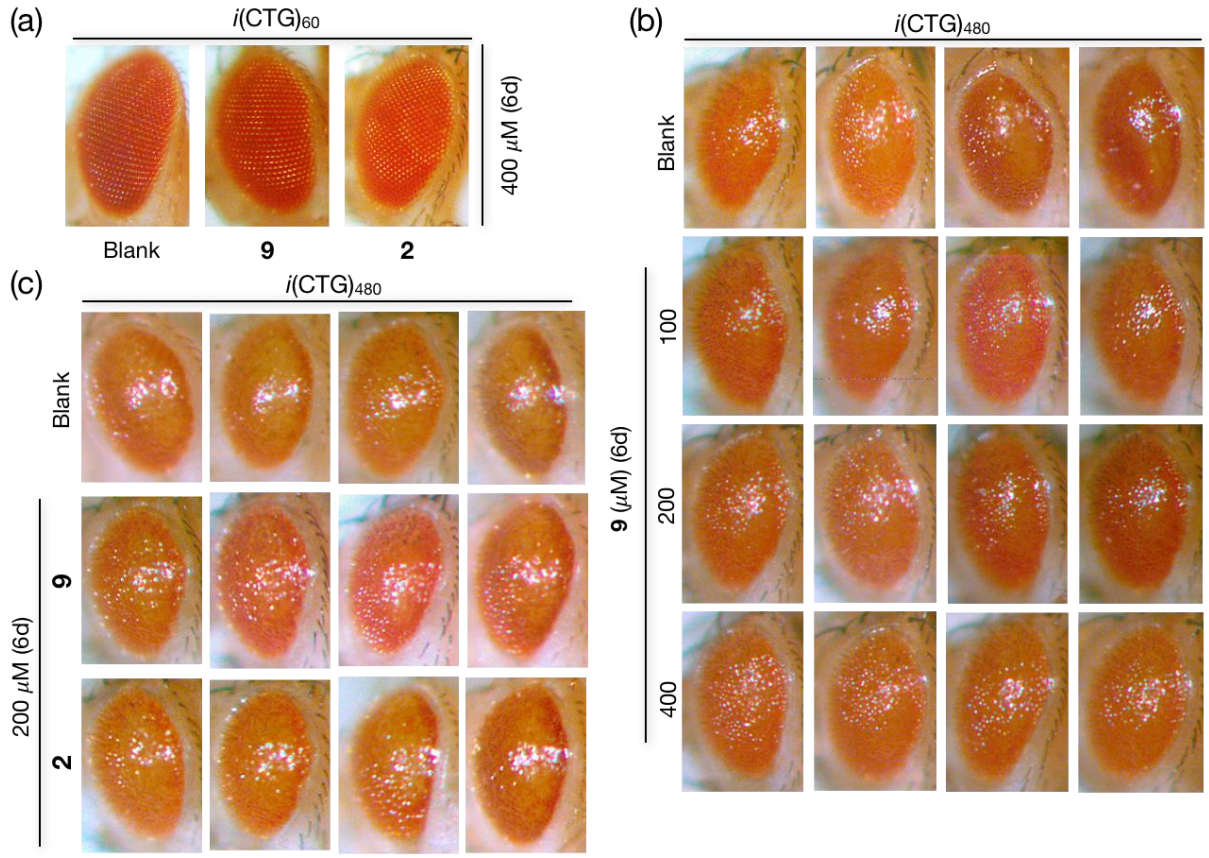


Figure 4.16. Ligands **2** and **9** rescue eye phenotypes in DM1 Drosophila. (a) Treatment of *Drosophila* bearing *i*$(CUG)_{60}$ with ligands **2** and **9** showed no effects on larval locomotion (left) and eyes (right), indicating that the ligands are relatively non-toxic. The error bars represent the standard deviation of three independent experiments (b) Ligands **2** and **9** reversed the disease phenotype in DM1 infected fly eyes at the concentration of 200 μM. (c) Ligand **9** rescued DM1 phenotypes in drosophila eye in a dose-dependent manner.

For these reasons, it was selected for *in vivo* testing in a DM1 *Drosophila* model,[12] specifically transgenic flies that express an interrupted $(CTG)_{480}$ sequence (*i*$(CTG)_{480}$). The experiments in the DM1 *Drosophila* model were performed by our collaborators, Edwin Chan and his research group at the Chinese University of Hong Kong. The flies exhibit severe neurodegeneration and manifest a number of disease symptoms, including the well-characterized glossy and rough-eye phenotype that can be easily observed by microscopy (Figure 4.16). Whereas ligands **2** and **9**

showed no effects on eye shape of *i*(CTG)$_{60}$ containing flies that do not have DM1 phenotypes, treatment of the DM1 flies with ligand **9** improved the neurodegenerative phenotype in a dose-

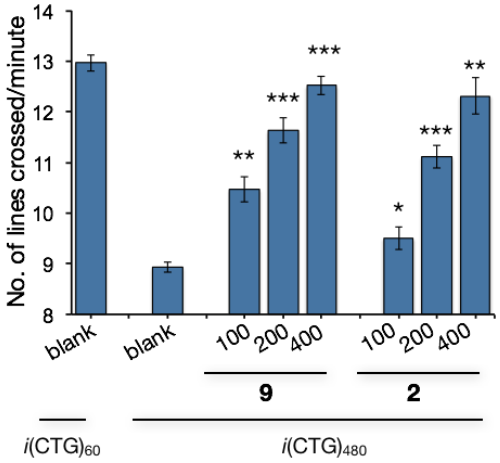


Figure 4.17. Larvae mobility was improved after the treatment of ligands. The error bars represent standard deviation of three independent experiments. For each independent experiment, 10 individual larvae were studied. *P < 0.05, **P < 0.01, ***P < 0.001 (two-tailed *t*-test).

dependent fashion. Particularly striking is a significant reduction in glossiness and a better-defined eye shape clearly observed after a 6-day treatment regimen (Figure 4.16). As we previously reported, control ligand **2** also showed reversal of the disease phenotype but less effectively.

These DM1 *Drosophila* exhibit other phenotypes, including impaired locomotion. To test the ability of **2** and **9** to improve the locomotor behavior of *Drosophila* larvae, we used crawling assays.[35-37] Untreated larvae having *i*(CUG)$_{60}$ do not show the phenotype and crawled with an average speed of ca. 13 lines/min, which is considered baseline locomotion. Those expressing *i*(CUG)$_{480}$ crossed only 9 lines/min (Figure 4.17b). Larvae that were treated with the highest doses of **2** and **9** (400 μM) showed significant improvement in locomotion with an average crawling speed approaching the normal baseline level of ca. 13 lines/min. Importantly, the phenotypic improvement with both compounds was dose-dependent, and **9**, in all cases, provided

greater phenotypic reversal than did **2**. Thus, with doses ranging from 100 to 400 μM, **9** exhibited between 38% and 89% recovery of normal locomotor behavior, whereas **2** showed between 14% and 83% recovery (Figure 4.17). Because **9** regulated the level of toxic CUG^{exp} in cells, we performed experiments to determine whether the same activity was observed in DM1 flies. *SV40* terminator mRNA is downstream from the $i(CUG)_{60}$ and $i(CUG)_{480}$ regions (Figure 4.18a). Thus, its amplification using specific primers that are detailed in the Experimental Section is directly correlated with CUG^{exp} mRNA levels. The *SV40* mRNA was expressed approximately equally in larvae bearing either $i(CUG)_{60}$ or $i(CUG)_{480}$ (Figure 4.18b). Treatment with **2** and **9** at 400 μM showed no change in *SV40* mRNA levels measured in larvae having

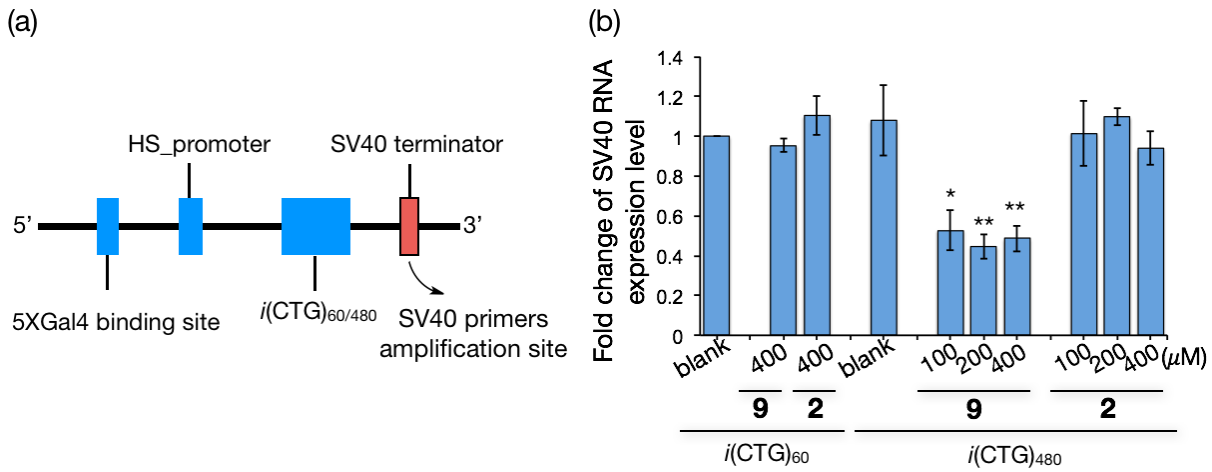


Figure 4.18. Ligand **9** regulates the level of toxic RNA in DM1 *Drosophila* model. (a) Diagram of $i(CTG)_{60/480}$ gene construct and the region (SV40 terminator) that was amplified to measure the level of transcribed mRNA. (b) Ligand 9 reduced the levels of *SV40* RNA in larvae. The *SV40* region of the CUG-containing RNA was measured relative to *β*-actin RNA level. The error bars represent the standard deviation of three independent experiments. For each independent experiment, mRNA levels from five individual larvae were determined; *P < 0.05, **P < 0.01 (two-tailed t-test).

$i(CUG)_{60}$. In contrast and consistent with the cell studies, **9** reduced by ca. 40−60% the *SV40* mRNA levels in the $i(CUG)_{480}$ larvae, whereas **2** did not. This result demonstrates the *in vivo* selectivity of **9** toward larvae expressing disease length CUG trinucleotide repeats.

4.8. Conclusion

The "holy grail" of DM1 therapeutic strategies would involve contraction of CTG^{exp} to nondisease lengths. Such a process could represent a cure for the disease. This is a particularly difficult challenge because multiple processes cause the expansion and their detailed mechanisms are not known. For this reason, drug discovery efforts to date have largely focused on the toxic CUG^{exp} transcript and its gain-of-function mechanism. A number of single-target small molecules are now known that selectively recognize CUG^{exp} and liberate sequestered MBNL1. However, recent reports indicate a more complex disease pathobiology, suggesting that binding CUG^{exp} may not be enough to reverse all disease pathways. Thus, efforts to destroy the toxic RNA transcript or inhibit its formation have particular appeal. Three reported examples of agents that control CUG^{exp} levels include antisense agents that induced CUG^{exp} cleavage via a RNase H-dependent manner,[22] a small molecule that degraded CUG^{exp} through a photo-induced cleavage,[24] and $(CTG{\cdot}CAG)_n$ transcription inhibitors.[25]

The efforts described herein sought a single small-molecule agent that might intervene in multiple DM1 disease pathways. We discovered agents **5**, **6**, and **9** that, indeed, operate in three distinct ways. Each shows RNase-A-like activity in selectively cleaving $(CUG)_{16}$ *in vitro*, and each inhibits both the *in vitro* transcription of CTG^{exp} and the sequestration of MBNL1 in nuclear foci in a DM1 model cell culture. Not all of the compounds performed equally well at each of these tasks. For example, **6** and **9** cleaved $(CUG)_{16}$ more rapidly than did **5**, and both fully inhibited nuclear foci formation, whereas **5** was not quite as effective. As a CTG^{exp} transcription inhibitor, **9** was the most effective but less selective than **6**. Interestingly, the least effective transcription inhibitor, **6**, was most effective at suppressing cellular levels of $(CUG)_{960}$. Although this might suggest that the RNA cleaving activity is most important, the cellular suppression of

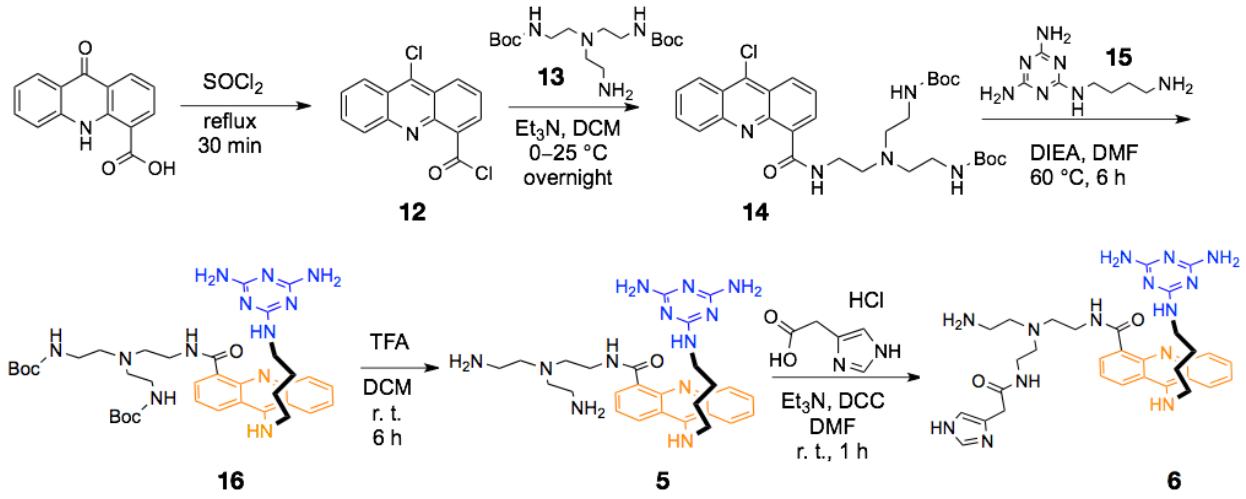
$(CUG)_{960}$ over the 3 day period appears to well out-pace the *in vitro* RNA cleavage rates. The latter are quite slow even at high compound concentrations. It is possible that one or more RNA nicks activate an endogenous RNase, or it may just reflect the complexity of the cell where $(CUG)_{960}$ suppression will depend on many factors, including cell permeability and the effectiveness of the three separate targeting activities in the complex environment of the cell. With all of the results and particularly the low cytotoxicity exhibited by **9** taken into account, it was considered the most promising candidate for *in vivo* studies. In a DM1 *Drosophila* model, **9** was found to rescue the neurodegeneration, thereby significantly reversing both the rough-eye phenotype and the larvae locomotor function. How does **9** function in the *Drosophila*? To the best of our knowledge, **9** is the first small molecule to control the level of CUG^{exp} in a DM1 model organism. It is beyond the scope of this investigation to determine how **9** functions *in vivo*, but the reduction in the toxic RNA levels supports the role of its multi-target ability observed in earlier studies. It was especially noteworthy that the control compound **2** shows no suppression of the CUG^{exp} levels.

None of the compounds described herein are able to recognize or affect the $(CAG)^{exp}$ transcript, so it is able to undergo RAN translation. It is also the case that agents such as **5**, **6**, and **9** that nick RNA and may act as transcription inhibitors will have to exhibit a very high level of selectivity to avoid undesirable off-target activity. Indeed, in the transcription assays with **9**, some inhibition was observed with the control sequences. Although achieving such selectivity remains as a future challenge, the repeating nature of the trinucleotide repeat target means that a bi- or polyvalent strategy can easily amplify selective targeting. Most significantly, we have demonstrated the first rational multi-target drug discovery effort that has led to three small molecules that intervene in three separate pathobiological steps in DM1, with one of the agents

showing enhanced phenotypic reversal in two separate DM1 *Drosophila* assays.

4.9. Materials and methods

Compounds, Materials, and General Methods. The preparation and characterization of ligands tested in the cellular experiments are reported herein. Unless otherwise noted, ^{1}H spectra were recorded on a 500 MHz Varian Unity Inova spectrometer. All NMR measurements were carried out at ambient temperature. Chemical shifts are in parts per million (ppm). Coupling constants (J) are reported in hertz (Hz). Electrospray ionization mass spectra (ESI-MS) were obtained by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois at Urbana-Champaign. All tested ligands are ≥95% pure as indicated by HPLC.



Scheme 4.1. Synthesis of ligands **5** and **6**

Compounds **11**, **12**, **13**, **15** were synthesized by following the reported procedures.[27]

Compound 14. To a 35 mL round-bottomed flask containing 500 mg (2.1 mmol) of compound 9-oxo-9,10-dihydroacridine-4-carboxylic acid (**11**) was added 5 mL (68.9 mmol) of $SOCl_2$ using a syringe and 4 drops of anhydrous DMF. The reaction mixture was refluxed for 1 h. The

unreacted $SOCl_2$ were removed azeotropically using DCM. The orange solid containing **12** was dried under high vacuum and used for the next reaction without further purification.

The crude product containing **12** was dissolved in 10 mL of anhydrous DCM and cooled in an ice bath. The pH of the solution was adjusted to pH = 11 using anhydrous Et_3N. A solution of 809 mg (2.3 mmol) of **13** in 5 mL of anhydrous DCM was added slowly to the reaction mixture. The reaction was slowly warmed to room temperature and stirred overnight. The reaction was monitored using TLC with a mixture of ethyl acetate:hexane = 8:2. Dichloromethane was removed using a rotary evaporator. The obtained orange solid was dissolved in 5 mL of ethyl acetate and purified by a silica column chromatography using a gradient mixture of ethyl acetate:hexane from 2:8 to 8:2, giving compound **14** as a yellow solid (280 mg, 23%). ^{1}H NMR (DMSO-d_6): δ 11.11 (br s, N*H*CO, 1H), 8.76 (dd, Ar*H*, 1H, *J* = 1.5, 5.5), 8.62 (dd, Ar*H*, 1H, *J* = 1.5, 7.5), 8.45 (d, Ar*H*, 1H, *J* = 7.5), 8.34 (d, Ar*H*, 1H, *J* = 9), 8.06-8.03 (m, Ar*H*, 1H), 7.90-7.85 (m, Ar*H*, 2H), 3.60 (q, CH_2, 2H, *J* = 5.5, 6), 3.08-3.02 (m, CH_2, 4H), 2.81 (t, CH_2, 2H, *J* = 6.3), 2.66 (t, CH_2, 4H, *J* = 6.3), 1.15 (br s, CH_3, 18H).

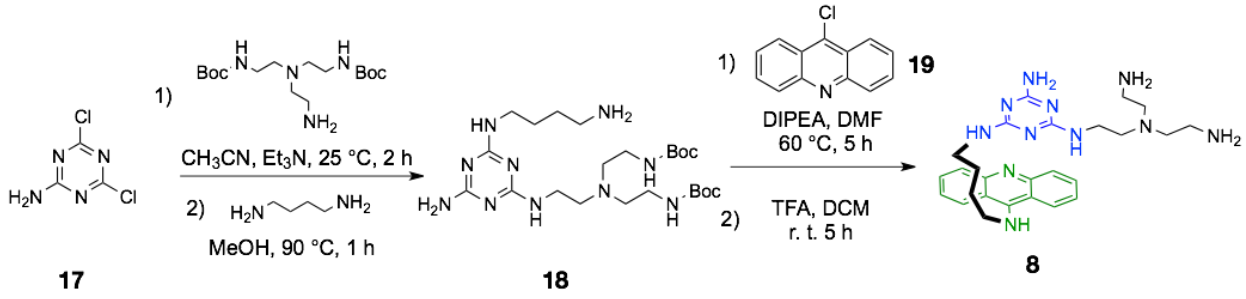
Compound 16. To a 50 mL round-bottomed flask containing 280 mg (0.5 mmol) of **14** in 20 mL of anhydrous DMF was added 109 mg (0.6 mmol) of N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine (**15**) and 0.19 mL (1.1 mmol) of DIPEA. The reaction was stirred at 60 °C for 6 h. The solvent was removed using vacuum, giving an orange solid. The crude solid was dissolved in 5 mL of DCM and purified by alumina column chromatography using a mixture of DCM:MeOH = 95:5, giving compound **16** as a orange solid (240 mg, 67%). ^{1}H NMR (DMSO-d_6): δ 8.60 (d, Ar*H*, 1H, *J* = 6), 8.55 (d, Ar*H*, 1H, *J* = 7), 8.42 (d, Ar*H*, 1H, *J* = 9), 7.97 (d, Ar*H*, 1H, *J* = 8.5), 7.75 (t, Ar*H*, 1H, *J* = 7.5), 7.53 (br t, N*H*CO, 1H, *J* = 7), 7.44-7.38 (m, Ar*H*, 2), 6.63 (br t, NH_2,

2, *J* = 5.5), 6.43 (br t, N*H*$_2$, 2, *J* = 5.8), 6.02 (br s, N*H*$_2$, 2), 5.87 (br s, N*H*$_2$, 2), 3.87 (br q, C*H*$_2$, 2H, *J* = 6.5), 3.54 (br q, C*H*$_2$, 2H, *J* = 6), 3.15 (q, C*H*$_2$, 2H, *J* = 6.5), 3.03 (br q, C*H*$_2$, 4H, *J* = 6.5), 2.76 (t, C*H*$_2$, 2H, *J* = 6.5), 2.65 (t, C*H*$_2$, 4H), 1.76 (br q, C*H*$_2$, 2H, *J* = 7.5), 1.51 (br q, C*H*$_2$, 2H, *J* = 7), 1.21 (br s, C*H*$_3$, 18H).

Ligand 5. To a 50 mL round-bottomed flask containing 280 mg (0.5 mmol) of di-tert-butyl (((2-(9-chloroacridine-4-carboxamido)-ethyl)azanediyl)bis(ethane-2,1-diyl))dicarbamate in 20 mL of anhydrous DMF were added 109 mg (0.6 mmol) of N2-(4-aminobutyl)- 1,3,5-triazine-2,4,6-triamine and 0.19 mL (1.1 mmol) of DIPEA. The reaction was stirred at 60 °C for 6 h. DMF was removed under high vacuum, affording an orange solid. The crude solid was dissolved in 5 mL of dichloromethane and purified by alumina (activated, basic) column chromatography using 95:5 (v/v) DCM/MeOH as eluent, giving 240 mg (67%) of the desired compound as an orange solid. To a 100 mL round-bottomed flask containing 240 mg (0.4 mmol) of the orange solid in 40 mL of dichloromethane was added 10 mL (131 mmol) of trifluoroacetic acid. The reaction was stirred at room temperature for 6 h. The solvent was removed using a rotary evaporator and dried under vacuum to give 410 mg (100%) of compound 5 as a yellow solid, TFA salt: ^{1}H NMR (D_2O) δ 8.40 (br d, Ar*H*, 1H, J = 8.5), 8.22 (br d, Ar*H*, 1H, J = 8.5), 8.12 (dd, Ar*H*, 1H, J = 1.5, 6), 7.81 (td, Ar*H*, 1H, J = 1.5, 7.5), 7.66 (d, Ar*H*, 1H, J = 7.5), 7.41 (br t, Ar*H*, 2H, J = 8.3), 4.05 (br t, C*H*$_2$, 2H, J = 6.5), 3.55 (t, C*H*$_2$, 2H, J = 7), 3.15 (br t, C*H*$_2$, 2H, J = 6.3), 3.10 (br t, C*H*$_2$, 4H, J = 6.5), 2.99 (br t, C*H*$_2$, 4H, J = 6.5), 2.93 (br t, C*H*$_2$, 2H, J = 7), 1.85 (br q, C*H*$_2$, 2H, J = 7), 1.58 (br q, C*H*$_2$, 2H, J = 7); ESI-MS (m/z) calcd for $[M + H]^+$ 547.3; found 547.3 $[M + H]^+$.

Ligand 6. To a 20 mL round-bottomed flask containing 200 mg (0.4 mmol) of compound **5** in 5

mL of DMF was added Et_3N to pH 7. The mixture of 101 mg (0.6 mmol) of 4-imidazole acetic acid and 129 mg (0.6 mmol) of DCC in 5 mL of DMF was added to the above solution. The reaction was stirred at room temperature overnight. The solvent was removed under vacuum. The crude product was purified using a reversed-phase C18 column on a CombiFlash system ($MeOH/H_2O$ (v/v) = 0:100 to 50:50) to afford 15 mg (12%) of compound 6 as a yellow solid (15 mg, 12%): ^1H NMR (D_2O) δ 8.56−8.50 (m, Ar*H*, 1H), 8.48 (br s, Ar*H*, 1H), 8.39−8.31 (m, Ar*H*, 1H,), 8.25 (br d, Ar*H*, 1H, J = 8), 7.94 (br t, Ar*H*, 1H, J = 8), 7.78 (br t, Ar*H*, 1H, J = 8.5), 7.54 (br t, Ar*H*, 2H, J = 8), 7.12 (s, C*H*, 1H), 4.19 (br t, CH_2, 2H, J = 6.8), 3.68 (br s, CH_2, 2H), 3.46 (br t, CH_2, 2H, J = 6.5), 3.30−3.21 (m, CH_2, 4H), 3.18−3.09 (m, CH_2, 4H), 3.02−2.96 (m, CH_2, 2H), 2.96−2.91 (m, CH_2, 2H), 2.03−1.96 (m, CH_2, 2H), 1.74−1.69 (m, CH_2, 2H); ESI-MS (m/z) calcd for $[M + H]^+$ 655.4; found 655.8 $[M + H]^+$, 328.5 $[M + 2H]^{2+}$.



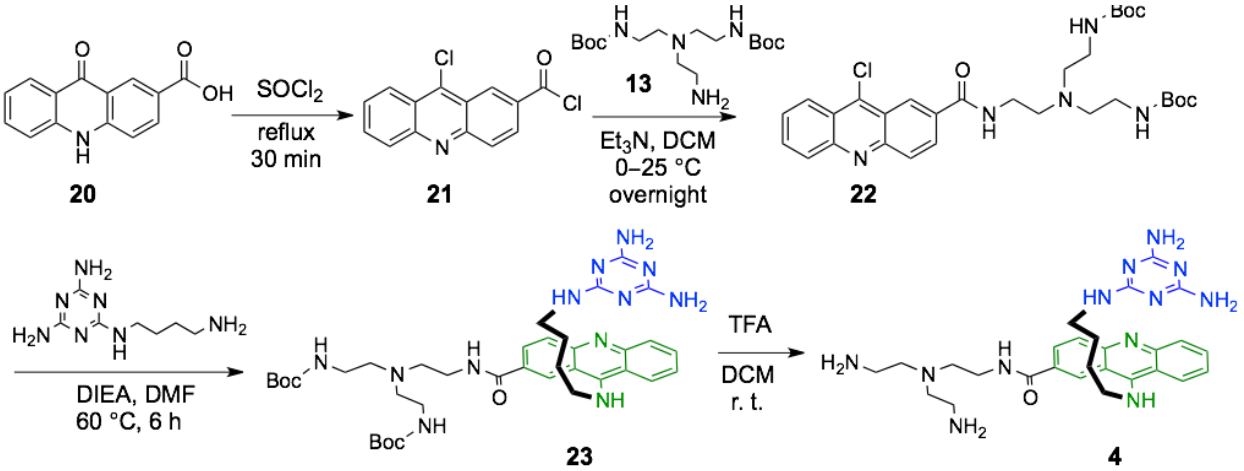
Scheme 4.2. Synthesis of ligand **8**

Compound 18. To a 50 mL round-bottom flask containing 690 mg (4.2 mmol) of compound **17**[1] was added 10 mL of acetonitrile and 0.7 mL (5.0 mmol) of triethylamine. The suspension was stirred at room temperature for 5 min. To the reaction flask was added 1.7 g (4.9 mmol) of **13** at once, giving a clear solution. The reaction was stirred at room temperature for 2 h. Acetonitrile was removed using a rotary evaporator. The obtained solid was dissolved in 30 mL of MeOH. To the resulting solution was added 2.0 mL (19.9 mmol) of 1,4-diaminobutane at once. The reaction

was heated to 90 °C, and stirred for 1 h. Methanol and the excess amount of diaminobutane were removed by a rotary evaporator. The crude was purified by silica gel column chromatography with a mixture of CH_2Cl_2/MeOH/NH_4OH = 90/9/1. Fractions containing product were combined and concentrated to give 1.3 g (59%) of compound **18** as a white solid. ^{1}H NMR (DMSO-d_6): δ 6.75 (br s, N*H*Boc, 2H), 6.53–5.81 (m, Ar*H*, 4H), 3.19–3.14 (m, CH_2, 4H), 2.97 (br s, CH_2, 4H), 2.56–2.53 (t, CH_2, 2H, J = 6.9), 2.46–2.44 (t, CH_2, 6H, J = 6.5), 1.49–1.30 (m, CH_2, 4H), 1.37 (br s, CH_3, 18H). ESI-MS (m/z) calculated for $[M + H]^+$: 527.4; found 527.2.

Ligand 8. To a 100 mL 24/40 round-bottom flask was added 850 mg (1.6 mmol) of melamine **18** and 420 mg (2.0 mmol) of acridine **19**. The mixture was dissolved in 20 mL of DMF. To the resulting yellow solution was added 0.3 mL (1.7 mmol) of DIPEA. The mixture was stirred at 60 °C for 5 h. The solvent was removed using a rotary evaporator. The crude was purified by alumina (activated, basic) column chromatography using a DCM/MeOH mixture (gradient from 100/0 to 97/3 (v/v)). The fractions containing product were combined and concentrated using a rotary evaporator. The yellow solid was dissolved in a mixture of 25 mL of TFA and 25 mL of DCM. The solution was stirred at room temperature for 5 h. The reaction was concentrated using a rotary evaporator, and dried overnight under high vacuum to give 876 mg (51%) of compound **8** as a yellow solid. ^{1}H-NMR (500 MHz, CD_3OD-d_4): δ 8.51−8.49 (m, 2H, Ar*H*), 7.96−7.94 (m, 2H, Ar*H*), 7.82−7.80 (m, Ar*H*, 2H), 7.55 (t, Ar*H*, 2H, J = 7.9), 4.22−4.21 (m, NHCH_2, 2H), 3.49−3.45 (m, NHCH_2, 4H), 3.10−3.05 (m, NH_2CH_2, 4H), 2.87−2.71 (m, NCH_2, 6H), 2.09−2.02 (m, CH_2CH_2, 2H), 1.82−1.74 (m, CH_2CH_2, 2H). ESI-MS (m/z) calculated for $[M + H]^+$: 504.3; found 504.3.

Compound 22. To a 35 mL round-bottom flask containing 300 mg (1.3 mmol) of **20** synthesized using the reported protocol[5] was added 2 mL (27.6 mmol) of $SOCl_2$ using a syringe and 8 drops of dry DMF. The reaction mixture was refluxed for 1 h. The unreacted $SOCl_2$ was removed azeotropically using DCM. The crude orange product containing **21** was dried under high vacuum and used for the next reaction without further purification.

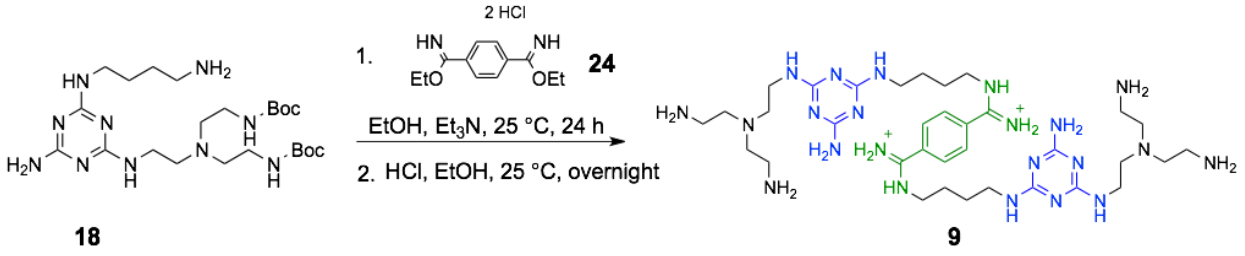


Scheme 4.3. Synthesis of ligand **4**

The crude product containing **21** was dissolved in 10 mL of anhydrous DCM and cooled in an ice bath. The pH of the solution was adjusted to pH = 11 using anhydrous Et_3N. A solution of 400 mg (1.2 mmol) of **13** in 5 mL of anhydrous DCM was added slowly to the reaction mixture. The reaction was slowly warmed up to room temperature and stirred overnight. The reaction was monitored using TLC with a mixture of ethyl acetate:hexane = 8:2. Dichloromethane was removed using a rotary evaporator. The resultant orange solid was dissolved in 5 mL of ethyl acetate and purified by silica column chromatography using a gradient mixture of ethyl acetate:hexane from 2:8 to 8:2, giving compound **22** as a yellow solid (300 mg, 25%). ^{1}H NMR (DMSO-d_6): δ 8.90 (t, Ar*H*, 1H, *J* = 3.3), 8.47 – 8.42 (m, Ar*H*, 1H), 8.29 – 8.24 (m, Ar*H*, 2H), 7.99 (ddd, Ar*H*, 1H, *J* = 8.4, 6.6, 1.4,), 7.82 (ddd, Ar*H*, 1H, *J* = 8.9, 6.6, 1.2), 6.69 (t, Ar*H*, 1H, *J*

= 5.7), 3.38 (q, CH_2, 2H, J = 6.5), 2.99 (q, CH_2, 4H, J = 6.5), 2.66 (t, CH_2, 2H, J = 6.8), 2.52 (t, CH_2, 4H, J = 6.8), 1.32 (br s, CH_3, 18H).

Ligand 4. To a 50 mL round bottom flask containing 270 mg (0.5 mmol) of compound **22** in 20 mL of anhydrous DMF was added 105 mg (0.5 mmoL) of N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine (**15**) and 0.184 mL (1.1 mmol) of DIPEA. The reaction was stirred at 60 °C for 6 h. DMF was removed using vacuum, affording an orange solid. The crude solid was dissolved in 5 mL of DCM and purified by alumina column chromatography with a mixture of DCM:MeOH = 95:5, giving compound **23** as a orange solid (260 mg, 72%). To a 100 mL round bottom flask containing 260 mg of compound **23** in 20 mL of DCM was added 10 mL of TFA. The reaction was stirred at room temperature for 6 h. The solvent was removed using a rotary evaporator and dried under vacuum to give compound **4** as a yellow TFA salt (100%). ^{1}H NMR (D_2O): 8.70 (t, ArH, 1H, J = 2.6), 8.22 (s, ArH, 1H), 8.04 (qd, ArH, 1H, J = 8.9, 1.7), 7.83–7.77 (m, ArH, 1H), 7.60 (t, 2H, ArH, J = 6.9), 7.40 (s, ArH, 1H), 4.13–4.07 (m, CH_2, 2H), 3.54 (t, CH_2, 2H, J = 6.9), 3.17 (q, CH_2, 2H, J = 1.7), 3.11 (br d, CH_2, 4H, J = 6.4), 3.04 (br s, CH_2, 4H), 2.96 (br s, CH_2, 2H), 1.88 (br s, CH_2, 2H), 1.60 (br s, CH_2, 2H).



Scheme 4.4. Synthesis of ligand **9**

Ligand 9. To a 100 mL oven-dried round-bottomed flask was added 250 mg (0.9 mmol) of diethyl terephthalimidate hydrochloride. The white solid was dissolved in 15 mL of anhydrous

ethanol. To the resulting suspension was added 0.3 mL (2.2 mmol) of anhydrous Et_3N followed by 1.0 g (1.9 mmol) of di-tert-butyl (((2-((4-amino-6-((4-aminobutyl)amino)-1,3,5-triazin-2-yl)amino)ethyl)azanediyl)bis-(ethane-2,1-diyl))dicarbamate at once. The resulting suspension was stirred at room temperature for 1 d. The solvent was removed using a rotary evaporator. The crude was dissolved in 20 mL of 2 N ethanolic HCl. The reaction was stirred at room temperature overnight. Ethanol was removed using a rotary evaporator. The crude was purified by a Sephadex CM-25 column chromatography using an aqueous solution of NH_4HCO_3 from 0.1 to 1.0 M. Fractions containing products were combined and concentrated at 60 °C using a rotary evaporator. The solid was dissolved in 80 mL of 0.1 M aqueous HCl. The resulting solution was concentrated using a rotary evaporator to give 390 mg (40%) of compound 9 as a white HCl salt: ^{1}H NMR (DMSO-d6) δ 10.27 (br s, N*H*, 2H), 9.80 (br s, N*H*, 2H), 9.46−9.41 (m, N*H*, 2H), 8.42−7.90 (m, Ar*H*, 8H), 8.00 (s, Ar*H*, 4H), 3.87 (br s, NH_2, 8H), 3.54−3.35 (m, CH_2, 12H), 2.96 (br s, CH_2, 8H), 2.76−2.62 (m, CH_2, 12H), 1.71−1.64 (m, CH_2, 8H); ESI-MS (m/z) calcd for $[M + H]^+$: 781.6; found 781.6.

RNA was purchased from Integrated DNA Technologies (Coralville, IA) and GE Dharmacon (Lafayette, CO). RNA samples were dissolved in THE Ambion RNA storage solution and stored at −20 °C. UV absorbance of the RNA solutions was measured at 25 °C on a Shimadzu UV-2501PC spectrophotometer. The concentration of the double-stranded RNA was calculated using Beer's law with the extinction coefficient at 260 nm provided by the supplier. The GFP-DT0 and GFP-DT960 plasmids were the gifts from Auinash Kalsotra (University of Illinois, Urbana-Champaign, IL); the DT960 minigene plasmid from Thomas Cooper (Baylor College of Medicine, Houston, Texas); the $(CTG)_{74}$ plasmid from Maurice Swanson (University of Florida,

Gainesville, FL, and the IR minigene plasmid from Nicholas Webster (University of California, San Diego).

RNA Cleavage Experiments. RNA was fast-folded at 95 °C for 5 min and then placed in ice for 10 min. RNA with a final concentration of 100 nM was incubated with 100 μM cleaving agents (in screening assays) or with ligand **9** at different concentrations (5, 10, 50, 100 μM). The cleaving buffer was 50 mM Tris buffer (pH 7.4) containing 150 mM NaCl and 2 mM $MgCl_2$. The final volume of the reaction mixture was 10 μL. The reaction was quenched by adding 8 μL of 8M urea and 2 μL of RNA loading dye or 1 μL of HIV FS RNA as a spike and 1 μL RNA loading dye (in screening experiments) followed by heating at 95 °C for 5 min. The reaction mixture was separated on a 20% RNA denaturing gel. For nonlabeled RNA, the gel was stained with EtBr and observed under UV. For TAMRA−$(CUG)_{16}$, the gel was scanned using a Typhoon instrument in the Biotech Center Lab (Noyes Laboratory, School of Chemical Sciences, University of Illinois, Urbana, IL). The images were worked up using ImageJ software (NIH). Foci dispersion and IR splicing experiments followed the reported protocols.[12]

Cellular mRNA Level Study with Dox Treatment. Approximately, 50 000 HeLa cells were plated on a 12-well plate in DMEM media supplemented with high glucose, L-glutamine, and no antibiotics a day before transfection. HeLa cells were transfected with 1 μg of GFP-DT0 or GFP-DT960 plasmids using Lipofectamine (Life Technologies) following the recommended protocol. After 4 h, the transfection cocktail was replaced with the growing media, and cells were treated with 1 μg of Dox. Ligands were treated at the same time at desired concentrations for 3 days. Cells were checked by fluorescence microscopy for a GFP signal as a marker of successful

transfection and then harvested. Total mRNA was isolated using E.Z.N.A. total RNA kit I (Omega). Approximately, 1.5 μg of total mRNA was subjected to DNase treatment to remove all DNA contaminant. cDNA synthesized using Iscript cDNA synthesis kit (Bio-Rad) was used as the template for real-time PCR using SYBR master mix (Applied Biosystem). The results from real-time PCR experiments were analyzed using the ΔΔCt method.[38] The mRNA levels of exon 15 upstream of CUGexp were measured relative to PABP mRNA levels. The difference in the expression level of exon 15 RNA transcript between treated GFP-DT0 and GFP-DT960 samples was compared with the one of untreated samples that were normalized to 100%. The primers were used in the experiments: E15upF: 5′-TCG GAG CGG TTG TGA ACT-3′; E15upR: 5′-GTT CGC CGT TGT TCT GTC-3′; PabpF: 5′-CTG CTG TTC ATG TGC AAG GT-3′; PabpR: 5′-CAA CAG CAT GCC AGT GAT T-3′.

In Vitro **Transcription of $(CTG{\cdot}CAG)_{74}$.** A 10 μL mixture contained 15 ng of linearized plasmid (CTG)74 or control plasmid templates, 0.5 mM each rATP, rCTP, rGTP, and rUTP, and 0.5 U T7 polymerase (Biolab) in 1X T7 transcription buffer (80 mM Tris pH 8.3, 10 mM $MgCl_2$, 2 mM spermine, 0.1% Triton-X, 10 mM NaCl) was incubated at 37 °C for 2 h. Ligand was added with the final concentrations of 0, 1, 10, 50, and 100 μM prior to incubation. Reactions were quenched by adding 8 μL of 8 M urea and 2 μL of denaturing dye (95% formamide, 5 mM EDTA, 0.025% each xylene cyanol and bromophenol blue) and heating to 95 °C for 5 min. Of this solution, 15 μL was run on a 8% denaturing polyacrylamide gel in 0.5X TBE. The gel was stained with EtBr. Bands were quantified using the ImageJ (NIH). The intensity of $(CUG)_{74}$ band (ca. 260 nt) was normalized to that from the untreated transcription reaction. The control plasmids were *pTRI-Xef* plasmid provided with MEGAscript T7 transcription kit (Life

Technologies) that expresses 1.89 kb RNA transcript and another non-repeat-containing plasmid expressing a 197 nt RNA transcript.

Drug Treatment in DM1 *Drosophila.* Drosophila lines were cultured in standard cornmeal medium supplemented with dry yeasts. Fly lines bearing UAS-(CTG)60 and UAS-(CTG)480[39,40] were kind gifts of Prof. Rubén Artero Allepuz (Universitat de Vale ncia, Estudi General, Spain). The gmr-GAL4[41] and 24B-GAL4[42] lines were used to drive UAS transgene expression in eye and muscles, respectively. Ligands 2 and 9 were dissolved in ddH2O and mixed with fly food. Genetic crosses were set up in drug-containing fly food at 21.5 °C for external eye assay and at 25 °C for larval crawling assay and real-time PCR analysis. For additional details, see REF. 12.

Larval Crawling Assay. Larval crawling assays were performed as described in Lanson et al.[35] Ten wandering third instar larvae were washed in ddH_2O and placed on a 2% agarose gel in a 15 cm Petri dish with gridlines spaced at 0.5 cm. The larvae were allowed to acclimate for a period of 1 min, and the total number of gridlines that the posterior end of the larvae passed in 1 min was determined. Each set of experiments was repeated independently three times using larvae collected from separate genetic crosses.

RNA Extraction and Real-Time PCR. RNA was extracted from third instar larvae by TRIzol reagent (Invitrogen). One microgram of purified RNA was used for reverse transcription via the ImPromII reverse transcription system (Promega). Real-time PCR gene expression assays were performed on an ABI 7500 real-time PCR system, using the SYBR Green PCR master mix (ABI) with the following primers: SV40_F: 5′-GGA AAG TCC TTG GGG TCT TC-3′;

SV40_R: 5′-GGA ACT GAT GAA TGG GAG CA-3′; actin_F: 5′-ATG TGC AAG GCC GGT TTC GC-3′ and actin_R: 5′-CGA CAC GCA GCT CAT TGT AG-3′. Each reaction was performed in duplicate. Quantification of gene expression was calculated according to the $2^{-\Delta\Delta Ct}$ method, where ΔΔCt = (Ct,target − Ct,actin)experimental − (Ct,target − Ct,actin)negative control. Each set of experiments was repeated independently three times using larvae collected from separate genetic crosses.

Isothermal Titration Calorimetry. ITC measurements were performed at 25 °C on a MicroCal VP-ITC (MicroCal, Inc., Northampton, MA). A standard experiment consisted of titrating 10 μL of a 500 μM ligand solution from a 250 μL syringe (rotating at 300 rpm) into a sample cell containing 1.42 mL of a 10 μM DNA or RNA solution. An ITC experiment consisted of 28 total injections (first injection was 5 μL, subsequent injections were 10 μL), with a 10 s duration per injection and delay of 380 s between injections. The initial delay prior to the first injection was 300 s. To derive the heat associated with each injection, the area under each isotherm (microcalories per second versus seconds) was determined by integration by the graphing program Origin 7.0 (MicroCal, Inc. Northampton, MA). The first data point from each ITC experiment was omitted when fitting to binding models because of the possible diffusive mixing of material near the tip of the syringe. The fitting requirements were such that the thermodynamic parameters were derived from curves that produced the lowest amount of deviation. In most cases, fitting to a sequential site-binding/model-binding site gave the most accurate data. The ligand stock solution was 10 mM in water. Double-stranded and hairpin DNA or RNA solutions were freshly prepared by mixing required volumes of the corresponding single-stranded oligomers and annealing by heating in a water bath at >90 °C for 5 min and

slowly cooling to room temperature. MOPS buffer solution (1 M), NaCl solution (5 M), and biological grade water were added to make up an oligonucleotide solution with 20 mM MOPS (pH 7.0 ± 0.2), 300 mM NaCl.

4.10. Reference

(1) Kunnumakkara, A. B.; Anand, P.; Aggarwal, B. B. *Cancer Lett.* **2008**, *269*, 199−225.

(2) Imming, P.; Sinning, C.; Meyer, A. *Nat. Rev. Drug Discovery* **2006**, 5, 821−834.

(3) Anighoro, A.; Bajorath, J.; Rastelli, G. *J. Med. Chem.* **2014**, *57*, 7874−7887.

(4) Morphy, R.; Rankovic, Z. *J. Med. Chem.* **2005**, *48*, 6523−6543.

(5) Meunier, B. *Acc. Chem. Res.* **2008**, *41*, 69−77.

(6) Gatchel, J. R.; Zoghbi, H. Y. *Nat. Rev. Genet.* **2005**, *6*, 743−755.

(7) Wheeler, T. M.; Sobczak, K.; Lueck, J. D.; Osborne, R. J.; Lin, X.; Dirksen, R. T.; Thornton, C. A. *Science* **2009**, *325*, 336−339.

(8) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. *J. Am. Chem. Soc.* **2008**, *130*, 16254−16261.

(9) Warf, M. B.; Nakamori, M.; Matthys, C. M.; Thornton, C. A.; Berglund, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 18551−18556.

(10) Parkesh, R.; Childs-Disney, J. L.; Nakamori, M.; Kumar, A.; Wang, E.; Wang, T.; Hoskins, J.; Tran, T.; Housman, D.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 4731−4742.

(11) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 16068−16073.

(12) Wong, C.-H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S.

C. *J. Am. Chem. Soc.* **2014**, *136*, 6355−6361.

(13) García-López, A.; Llamusí, B.; Orzáez, M.; Pérez-Payá, E.; Artero, R. D. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 1*08*, 11866−11871.

(14) Jahromi, A. H.; Honda, M.; Zimmerman, S. C.; Spies, M. *Nucleic Acids Res.* **2013**, *41*, 6687−6697.

(15) Cho, D. H.; Thienes, C. P.; Mahoney, S. E.; Analau, E.; Filippova, G. N.; Tapscott, S. *J. Mol. Cell* **2005**, *20*, 483−489.

(16) Moseley, M. L.; Zu, T.; Ikeda, Y.; Gao, W.; Mosemiller, A. K.; Daughters, R. S.; Chen, G.; Weatherspoon, M. R.; Clark, H. B.; Ebner, T. J.; Day, J. W.; Ranum, L. P. W. *Nat. Genet.* **2006**, *38*, 758−769.

(17) Zu, T.; Gibbens, B.; Doty, N. S.; Gomes-Pereira, M.; Huguet,A.; Stone, M. D.; Margolis, J.; Peterson, M.; Markowski, T. W.; Ingram, M. A. C.; Nan, Z.; Forster, C.; Low, W. C.; Schoser, B.; Somia, N. V.; Clark, H. B.; Schmechel, S.; Bitterman, P. B.; Gourdon, G.; Swanson, M. S.; Moseley, M.; Ranum, L. P. W. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 260−265.

(18) Wojciechowska, M.; Olejniczak, M.; Galka-Marciniak, P.; Jazurek, M.; Krzyzosiak, W. J. *Nucleic Acids Res*. **2014**, *42*, 11849−11864.

(19) Kalsotra, A.; Singh, R. K.; Gurha, P.; Ward, A. J.; Creighton, C. J.; Cooper, T. A. *Cell Rep.* **2014**, *6*, 336−345.

(20) Laurent, F. X.; Sureau, A.; Klein, A. F.; Trouslard, F.; Gasnier, E.; Furling, D.; Marie, J. *Nucleic Acids Res.* **2012**, *40*, 3159−3171.

(21) Pettersson, O. J.; Aagaard, L.; Andrejeva, D.; Thomsen, R.; Jensen, T. G.; Damgaard, C. K. *Nucleic Acids Res.* **2014**, *42*, 7186−7200.

(22) Lee, J. E.; Bennett, C. F.; Cooper, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*,

4221−4226.

(23) Wheeler, T. M.; Leger, A. J.; Pandey, S. K.; MacLeod, A. R.; Nakamori, M.; Cheng, S. H.; Wentworth, B. M.; Bennett, C. F.; Thornton, C. A. *Nature* **2012**, *488*, 111−115.

(24) Guan, L.; Disney, M. D. *Angew. Chem., Int. Ed.* **2013**, *52*, 1462−1465.

(25) Coonrod, L. A.; Nakamori, M.; Wang, W.; Carrell, S.; Hilton, C. L.; Bodner, M. J.; Siboni, R. B.; Docter, A. G.; Haley, M. M.; Thornton, C. A.; Berglund, J. A. *ACS Chem. Biol.* **2013**, *8*, 2528−2537.

(26) Wong, C.-H.; Richardson, S. L.; Ho, Y.-J.; Lucas, A. M. H.; Tuccinardi, T.; Baranger, A. M.; Zimmerman, S. C. *ChemBioChem* **2012**, *13*, 2505−2509.

(27) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. *ACS Chem. Biol.* **2013**, *8*, 1037−1043.

(28) Palmer, A. J.; Wallace, H. M. *Amino Acids* **2010**, *38*, 415−422.

(29) Trawick, B. N.; Daniher, A. T.; Bashkin, J. K. *Chem. Rev.* **1998**, *98*, 939−960.

(30) Lönnberg, H. *Org. Biomol. Chem.* **2011**, *9*, 1687−1703.

(31) Breslow, R.; Huang, D. L.; Anslyn, E. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 1746−1750.

(32) Miller, J. W.; Urbinati, C. R.; Teng-Umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. *EMBO J.* **2000**, *19*, 4439−4448.

(33) Wong, C. H.; Fu, Y.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Nucleic Acids Res.* **2011**, *39*, 8881−8890.

(34) Savkur, R. S.; Philips, A. V.; Cooper, T. A. *Nat. Genet.* **2001**, *29*, 40−47.

(35) Lanson, N. A.; Maltare, A.; King, H.; Smith, R.; Kim, J. H.; Taylor, J. P.; Lloyd, T. E.; Pandey, U. B. *Hum. Mol. Genet.* **2011**, *20*, 2510−2523.

(36) Nichols, C. D.; Becnel, J.; Pandey, U. B. *J. Visualized Exp.* **2012**, *61*, e3795.

(37) Batlevi, Y.; Martin, D. N.; Pandey, U. B.; Simon, C. R.; Powers, C. M.; Taylor, J. P.; Baehrecke, E. H. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 742−747.

(38) Schmittgen, T. D.; Livak, K. J. *Nat. Protoc.* **2008**, *3*, 1101−1108.

(39) Garcia-Lopez, A.; Monferrer, L.; Garcia-Alcover, I. *PLoS One* **2008**, *3*, e1595.

(40) Garcia-Alcover, I.; Colonques-Bellmunt, J.; Garijo, R.; Tormo, J. R.; Artero, R.; Álvarez-Abril, M. C.; López Castel, A.; Pérez-Alonso, M. *Dis. Models & Mech.* **2014**, *7*, 1297−1306.

(41) Freeman, M. *Cell* **1996**, *87*, 651−660.

(42) Brand, A. H.; Perrimon, N. *Development* **1993**, *118*, 401−415.

Appendix

Publication

(1) **Nguyen, L**.; Luu, M. L.; Peng, S.; Serrano, F. J; Chan, W.-Y.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2015**, *137*, 14180-14189.

(2) Luu, M. L; **Nguyen, L.;** Chan, W.-Y.; Zimmerman, S. C. Dimeric Bisamidinium-based Ligand as a Strong Inhibitor for MBNL1-CUG^{exp}, **2015**, accepted (co-first author)

(3) **Nguyen, L.**; Lee, J.; Wong, C.-H.; Zimmerman, S. C. *ChemMedChem* **2014**, *9*, 2455-2462. (designated VIP, highlighted on the journal cover)

(4) Wong, C.-H.; **Nguyen, L**.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Ho, T.; Chan, E. H. Y.; Chan, W.-Y.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 6355−6361.

(5) Jahromi, A. H.; Fu, Y.; Miller, K. A.; **Nguyen, L.;** Luu M. L.; Baranger, A. M.; Zimmerman, S. C. *J. Med. Chem.* **2013**, *56*, 9471–9481. (co-second author)

(6) Jahromi, A. H.; **Nguyen, L**.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. *ACS Chem. Biol.* **2013**, *8*, 1037–1043. (highlighted on the journal cover and the department news)